RESEARCH ON UTILIZATION OF PART TASK SPATIAL ORIENTATION INFORMATION IN THE DYNAMIC SIMULATOR

TECHNICAL REPORT NO. 299-099-284

NASA CONTRACT NASw-439

JUNE 1965

N 66-12195
(ACCESSION NUMBER) (THRU)
86 1
(PAGES) (CODE)
CR 68219 05
(NASA CR OR TMX OR AD NUMBER) (CATEGORY)
FACILITY FORM 602

GPO PRICE $
CFSTI PRICE(S) $
Hard copy (HC) 3.00
Microfiche (MF) .75
ff 653 July 65

BELL HELICOPTER COMPANY

FORT WORTH, TEXAS

DIVISION OF BELL AEROSPACE CORPORATION • A textron COMPANY

BELL HELICOPTER COMPANY

Technical Report

No. 299-099-284

RESEARCH ON UTILIZATION OF PART TASK SPATIAL ORIENTATION INFORMATION IN THE DYNAMIC SIMULATOR

June 1965

by: C. B. Elam
C. B. Elam
Director of Human Factors
Research

Billie A. Abbott
Billie A. Abbott
Human Factors Engineer

APPROVED:

D. J. Dougherty
D. J. Dougherty
Chief, Human Factors

H. W. Mitchell
H. W. Mitchell
Chief Electronics Engineer

National Aeronautics and Space Administration
Contract NASw-439

This report is submitted in compliance with the contractual obligation outlined in Contract NASw-439 between the National Aeronautics and Space Administration and Bell Helicopter Company.

The title, "Research on Utilization of Part Task Spatial Orientation Information in the Dynamic Simulator", which identified the initial proposal has also been used for this report.

ABSTRACT

12195

Fifteen professional pilots and a like number of R.O.T.C. students were used as subjects in establishing the relative difficulty in controlling for pitch, roll, yaw and altitude when these parameters were presented in all possible combinations with one another. The Bell simulator was used with the Norden vertical display serving as the media for information transmission. Momentary error and control position were recorded and were later converted into error and inefficiency scores.

While highly significant in a statistical sense, the differences between conditions were considered to be of limited theroretical interest. The primary contribution of the study was in its comparison of the different indices of error (absolute, squared and standard deviation) and the measures of inefficiency.

It was concluded that the standard deviation is the most sensitive of the error measurements and that an index of inefficiency (product of the error and the rate of control movement) is more sensitive of error alone.

Author

ACKNOWLEDGMENTS

The authors wish to express their endebtedness to Mr. Robert Taylor and Mr. Ray Bohling of the National Aeronautics and Space Administration for their recommendations in the design of the experiment and their patience in the development of the data.

A particular acknowledgment of gratitude is due Mr. Max Driggers and Mr. Ralph Absolom who maintained the simulator and its associated equipment during the extended period of data collection. Finally, this list would be incomplete without the mention of Mr. Leonard Lindenmeyer who presided over the data conversion.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF TABLES

I. INTRODUCTION

This study presumes to examine two general research areas. First, its intent is to describe tracking behavior as a consequence of task complexity where complexity relates to the number of subtasks making up the composite effort. Second, various analytical procedures are examined with the purpose of determining their relative value to the description of tracking performance.

Historically, the systematic study of tracking performance begins on a theoretical rather than on an applied level. The procedures and methods of measurement developed in this early work continue to exert some influence on contemporary thinking, although the emphasis has greatly altered. The original work was performed in an effort to describe human behavior and capabilities in a very general way. More recently, the trend has been toward a description of machines as they are operated upon by the human. Comparison is made of systems by utilizing the human operator as the criteria for evaluation or judgment. This approach has become more pronounced with the development of simulators which are used for training and experimental purposes. These devices can be used either for very practical and specific purposes or as tools for theoretical research. This versatility has produced some confusion in their use. When the experimental effort is primarily related to questions of human ability, a somewhat different approach is required than when the questions pertain to system utilization. Some measurements, for example, will be appropriate to one type of research, but of dubious value to the other.

In the operation of a particular aircraft, it is reasonable to inquire into the disadvantages of giving the pilot tasks to perform in addition to those for which he is already responsible. In such a study, the independent variables can be specified and suitable measures of performance can be contrived.

As a theoretical study (one in which the data can be extrapolated upon in application to a number of dissimilar situations), however, the question of performance as a function of task loading is difficult to answer since it is impossible to operationally define the term "task" without narrowing it to the point where it no longer retains any general value.

Added to this is the fact that the measure of performance is not a very objective way of describing human behavior since by its very nature, it sets up an external and possibly arbitrary criteria.

In retrospect, the present study does not appear to have avoided these difficulties. Its principle merit perhaps lies in the fact that it brings some of these problems into sharper focus.

II. STATEMENT OF THE PROBLEM

There is a considerable amount of literature in the area of control as it is affected by the amount, the complexity and the perceptual unity of the sub-tasks involved. These studies have often been categorized into such research areas as work load, part task versus whole task training, transfer of training, etc. Despite the obvious commonality of these subjects, it is difficult to take a systematic account of the results. The reason for this is due, to some extent, to the fact that many different types of tasks have been involved. Some made use of discrete and discontinuous events like throwing a switch or opening a valve. Others required continuous manipulation such as tracking the oscillations of a signal on a CRT. Still others require verbal responses. Often the studies make simultaneous involvement of all these elements.

Added to this heterogeneity is the fact that there is little agreement on what constitutes a description of performance. Some workers report errors while others analyze time to respond or time on target. The possibilities for measurement or evaluation are apparently infinite, being limited only by the experimenter's judgment of relevance. This is, of course, understandable since it is the nature of tasks to differ greatly. The criteria for evaluation will also differ if the work is to have any applicability to the problem areas selected.

The present investigation does not presume to settle these questions and certainly no criticism is directed toward any previous work. It does, however, attempt to bring the problem into somewhat better perspective by considering alternative methods of measurement.

The first consideration, and perhaps the one for which there is no satisfactory theoretical basis for settlement, is simply a question of definition. What is the proper unit for the description of a task? In common usage, and indeed in scientific description, the application of the term is altogether arbitrary. It may consist of an eyeblink, driving a car or being president of a corporation. We can speak of piloting an aircraft as a task or we can analyze it into two tasks: one involving flight control and the other engine control. Each of these can undergo further fragmentation or analysis. Still other divisions can be made such as considering navigation and aircraft control to be separate tasks. It is evident that all of these divisions are abstract, being created for analytical convenience.

This discussion is to engage in something other than the controversy of whether the basic units of behavior are to be described in molar or molecular terms. Whatever can be said for the impeccability of its logic, it should be obvious that a description of piloting behavior based upon muscle twitches would be impractical. The question then is not whether the units of task description should involve complex and heterogeneous elements, since it is evident that they must, but rather it must concern itself with the selection of operationally significant criteria which can be applied to those collections we employ. The difficulty transcends the problem of communicating with people who read such reports as this. The subjects of an experiment are also caught up into the artificialities of its analysis. The subject may be told to hold a symbol in the center of a CRT display by moving a control stick laterally when the symbol makes sideways excursions and to push it fore and aft when it moves up or down. The experimenter would almost unconsciously fall into this manner of speaking if he intends to analyze the data in terms of an X-Y coordinate system. Yet X and Y coordinates are not the perceptual endowment of all humans. Indeed they are innate to none.

On the other hand, the experimenter can simply tell the subject to keep the symbol in the center of the scope by moving the lever in a direction he will subsequently find to be appropriate.

Each of these instructions are to some degree prejudicial to experimental control. The first almost certainly implies to the subject that two tasks are involved. The second suggests that a unitary approach is to be taken. Either or both may be inappropriate when analysis is made.

One is almost tempted to recommend that perceptual unities be employed as the indivisible quanta of motor performance but this, too, is not without its difficulties. By perceptual unity, one would mean those tasks the subject sees as being necessary and sufficient for the control of a parameter. Thus in flying a helicopter, those aspects most closely associated with the cyclic (pitch and roll) would perhaps be perceived in some unified context while altitude (which depends primarily upon the collective stick) would have an independent status. The problem here, of course, is that subjects would differ between themselves and even the same subject would change his conceptual organization of the total task as the experiment progressed.

The reason for this preoccupation with the problem of task identification should be evident. To many workers, it seems reasonable to inquire into the relationship between performance and work load. The problem arises, however, in trying to give quantitative expression to work load in a manner that is independent of the performance measurement. This is to say that the independent variable must be defined differently than the dependent variable is. To Task A is added Task B. Aside from any measure of performance, it is doubtful if the work load has been exactly doubled since the tasks may be separately unequal. Some coefficient of adjustment could conceivably be made for these differences if it were true that in reality the work load was additive, despite the different scales of measurement. But it is almost certain that the additive principle does not always apply. Not only is the composite of A and B not twice as great in work load than A or B but there are instances where the aggregate of A and B produce less error than

A or B taken separately. Some tasks integrate with others to the advantage of each. Others produce conflicts which yield unexpected accumulations of error.

In trying to give graphic description to the performance-work load function, the experimenter has an ordinate (performance) but appears to have an abscissa (work load as measured by task analysis) whose points may migrate and at times transpose themselves. In any case, they do not lend themselves to linear scaling.

This being the case, it is with some dismay that we recognize that our ordinate is perhaps even less respectable than our abscissa.

Up to this point this scale has been referred to rather loosely as representing performance in a motor task. Usually, perfect performance is thought of as the absence of error. This being true, one can convert error scores into performance scores by taking reciprocals or by subtracting them from some realistic constant. Thus, they are only as good as our concept of error allows them to be.

Much has been made of the value of objective research, particularly quantitative applications. It is almost as though we were convinced that arithmetic operations, through some mystical transmutation, can impose order on what is truly chaotic. Not so many years ago, and for many of those in the area of human engineering, this has a disturbing freshness, all of this was handled in a simpler fashion. Usually, if one had a system to evaluate an expert, a pilot for example, was asked to use it. Later he was asked for his opinion. If his answer was complimentary, this was regarded as sufficient reason for further development or promulgation of the design. If his reply was negative, the experimenter had the option of changing the design or changing the pilot.

It was inevitable that this method should come under some criticism since it was difficult to keep individual bias from intruding into the evaluation. Added to this was the fact that there was no real way of determining how the judgments were being made since obviously they were highly subjective affairs. Granting the questionable nature of this procedure, it should be recognized that the pilot probably did not commit some of the gross atrocities that are perpetrated in the name of objective quantitative research. The pilot at least recognized that the importance of error is a highly relative thing. In some subjective fashion, he probably integrated what he perceived to be his errors into the production of an overall judgment, but it is doubtful that he did this in a blind mechanical way. He knew, for example, that some errors are more important than others and that a given error can be tolerated under one circumstance that would be fatal under another. He knew to expect more errors when he was relaxed than when he was working hard.

There are many other such variables which the pilot might make a proper interpretation of, but which the experimenter, for reasons of convenience of analysis, would choose to ignore. The latter might simply accumulate absolute error without reference to the momentary situation. This might be a proper thing to do in the context of pure research although even then it is doubtful. It is of most dubious validity when applied to the use of flight vehicles. In flying a terrain following mission, for example, the pilot has two things to consider. First, of course, he must avoid obstacles. The probability of this becomes increasingly great as he flies higher. At the same time, he must exercise caution about flying too high since there is the risk of being detected by enemy surveillance. Taken together, these two conflicting variables produce a non-linear probability function having a saddle point of optimal safety. An analysis based upon absolute error from command is obviously inappropriate to apply to the data; yet, this is frequently done.

It is convenient to regard uncontrolled errors as belonging to three classes. First are those which the operator recognizes to exist but to which he does not respond to because he feels that they are unimportant. Second are those errors which the operator recognizes to exist, but which, due to his own ineptitude or to deficiencies in the system, he is unable to null out. Third are those errors which the operator does not recognize to exist and consequently does nothing to eliminate.

Presuming that a procedure can be found to distinguish between them, these error categories should be handled differently. This, in part, is the objective of the present study.

Although conducted in a helicopter simulator, this study was regarded as fundamentally theoretical. This being so, it was not feasible to adopt an error scale that conformed to the particular problems of any given vehicle or system. Recognizing that any scale used for the assessment of error would, in the final analysis, be arbitrary, it was decided to investigate several scales with the view of obtaining one which would best conform to the subject's evaluation of importance. The three selected were absolute error, squared error and the standard deviation of error. The rationale was based on the idea that the scale which corresponded most closely to the subject's evaluation of importance should produce the greatest differences (as measured by the F ratio of an analysis of variance) between test conditions if these test conditions were truly of differential difficulty.

A second procedure introduced here was contrived with the intention of measuring subject efficiency. This relates to the question of whether the subject regards the momentary error as being important enough to respond to.

The general notion of efficiency can be expressed as the ratio between error and magnitude of activity. If an error exists but the subject does nothing to null it out, he is not behaving inefficiently although he may be behaving inappropriately from the experimenter's point of view. On the other hand, if an error exists which requires an unusual amount of activity from the subject in order to correct it, his behavior can be characterized as inefficient. Thus, by taking the product of error and a measure of the subject's activity, one can produce an index of inefficiency. In this study, the measurement of activity was simply the integrated rate of control movement. Arithmetical products were obtained between each of the measurements of error described above and the measurement of the work performed. This in turn produced three measures of inefficiency. These measures are described more precisely in a subsequent section of this report.

III. BACKGROUND

The early history of experimental psychology contains numerous descriptions of studies of performance on simple manual tasks. Much of this is systematic and of continuing theoretical interest. The measurements were often of reaction time to the onset of lights or other stimuli with the view of taking measurements of sensory, motor and cognitive processes. Some work was accomplished on the effect of two or more concurrent tasks on performance and, while the experimental arrangements were not specific to any particular control problem, this work also retains some general interest. Except as these effects tended toward individual differences, they did not lend themselves to the generation of many useful concepts or theories, however.

As with many other areas in scientific investigation, World War II served an an impetus to research on performance in complex task situations. In the past few decades a further acceleration of interest on manual control is evident. This has led to multiple bifurcations of subject matter which generally have acquired independent linguistic developments. It would not be appropriate to trace these mutations here. There are, however, three general areas which can be said to relate to the present study. These are: (1) the effect of varying work load on psychomotor performance; (2) part task vs whole task training; and (3) transfer of training.

The work of Conrad (2.) is germane to the problem of work load. In this study, the effect of an increase in perceptual load on performance was examined. Subjects were given the number of dials and an equal number of control handles which were in paired correspondence with one another. The task of the subject was to monitor the dials and to correct, by means of the appropriate control handle, any motion away from the standard that might occur. Beginning with four dials, Conrad successively increased the task until 12 dials required monitoring and control. If movement occurred on a dial, it would continue unchecked. If its movements were observed, the needle could be returned to its initial position by means of the handle. Scoring was obtained of the number of corrections occurring each minute and of the elapsed time between the initiation of a movement and the correcting response. It was found that as the number of tasks and dials to be monitored and controlled increased, there was a proportionate decrease in the number of corrections per minute and of corresponding increase in the elapsed time before correction was made. It was, however, suggested that a plateau was reached with about 10 dials because no significant difference was obtained between the 10 and the 12 dial conditions. Although Conrad terminated his investigations with 12 dials, it is possible that the leveling-off process would have continued beyond this number. Jeantheau (6.) continued this line of investigation by testing the differential effects of speed and load stress on task performance. He investigated the independent effects of variations in task speed and load in an information processing task. Load was varied in terms of word length with alphabet size and words per measure held constant. Speed was varied as the number of words per minute. The results indicate that both speed and work load increases impair performance. Significant, however, is the fact that the shortest word length had little effect on performance as the rate of presentation increased. This appears to suggest that for tasks of a simple nature, there is little decrement in performance with increased load, but that this generalization does not apply to the more complex tasks.

Support for this conclusion was obtained by Baker, et al (1.). These investigators studied the effects of complex additional tasks on performance. They obtained a continuous impairment on performance as the task loading was increased. If the leveling-off phenomenon occurs, it would then be limited to simple tasks.

The difficulty in making comparisons of this kind lies in the definition of simple and complex, as they relate within the present context. Conceivably these could merely be an unfortunate use of terms. Simple may be better expressed in description of tasks which are easily integrated because of the similarity or conceptual unity. Complexity then would refer to dissimilar tasks, especially those in which transfer of training is either absent or negative. This is only to restate the problem introduced in the previous section. As work is added, are the new tasks simply multiple reiterations of the original tasks or are they truly independent?

The classical work on transfer of training is too familiar to require description here. The reader is referred to McGeoch (5.) for an extensive treatment of this topic. In a recent study by Elam and Emery (4.), the data appeared to lend credence to two hypotheses. The first was that, within limits, an increased work load, made up of tasks showing positive transfer to one another, produced a negatively accelerated error function, while tasks showing negative transfer produced a positively accelerated growth function. The second hypothesis was that the measurement of efficiency (ratio of error to work performed) produces a superior analysis of task difficulty than error scores alone. This is based upon the notion that subject error as distinguished from total system error exists only if the subject acknowledges it to exist by making proper control movement.

IV. METHOD

Subjects

Thirty subjects were used in the collection of the data of this experiment. Fifteen of these were rated helicopter pilots obtained from the Flight Test Department of Bell Helicopter. The remaining subjects were without flight control experience. They were selected from the ROTC program of Texas Christian University. All subjects were paid for their participation in this study. It was ascertained for all subjects that vision (corrected) was normal and that none suffered from other debilitating sensory or motor impairment. The choice of the subject population was based upon the desire to obtain groups that were roughly similar to those individuals who might be used in other NASA studies. It was assumed that since both the ROTC and the pilot group had previously satisfied certain physical and mental criteria, they would be relatively homogeneous as compared with the general population. This statement is especially true for the pilots, although it is believed that the data taken from the performance of the non-pilot subjects can also be extrapolated upon to those operator situations that are of particular interest to NASA. The pilot and the non-pilot subject groups were treated alike except that while the former experienced each experimental condition only once, the non-pilots repeated the sessions on 15 separate occasions. This procedure would have been followed for the pilots but it was impossible to schedule them for so sustained a period. In consequence, the pilot group may be regarded as a replication of the non-pilots for the first day of testing of this latter group. Although the pilots were familiar with the control situation as found in the flight simulator, it should be pointed out that they were not familiar with the display used in this study. Added to this was the fact that the simulator did not respond in a manner that these subjects were accustomed to obtaining in an aircraft. One particularly novel element was that the simulator did not translate fore, aft or laterally as a flight vehicle would do. Nor was there any cross-coupling between control elements as is found in actual flight vehicles. These peculiarities, in concert with the novelty of the display media, were undoubtedly disconcerting to the pilots. These remarks are entered as a caution to the reader not to expect the pilot performance to be of a consistently high quality.

Apparatus

Simulator: All testing was accomplished with the subjects in the Bell Dynamic simulator. This apparatus was constructed in simulation of the side-by-side cockpit control configuration. The controls, standard for helicopter, consisted of cyclic stick for roll and pitch control, collective stick for altitude control and foot pedals for yaw control. Although the moving platform upon which the cabin rested was capable of six characteristics of motion (roll, pitch, yaw, altitude, lateral and horizontal displacement) only four (roll, pitch, yaw and altitude displacement) were actualized in the present experiment. The system dynamics, in a very gross manner, were memetic of helicopter flight--with the exception that all cross-coupling was removed and neither the display nor the platform was free to exhibit either lateral or horizontal translation. While it was recognized that these restrictions greatly reduced the similarity between the apparatus and any flight vehicle, it should be understood that these modifications were essential to the conduct of the experiment. Although it was desirable to preserve as much similitude with an actual flying situation as possible, this factor was of secondary importance. Since the purpose of the study was to examine performance as a consequence of the number of unitary tasks involved, it then became essential that these tasks remain absolutely independent to one another. Thus, the presence of cross-coupling would have been confounding to the analysis since there would be a multiplicity of control possibilities; none of which would absolutely enjoy a preferential status. If altitude could be affected directly through the collective or indirectly through the pitch channel, it would be impossible to relate error to input in any systematic manner. The problem, moreover, would be intensified under those conditions where a particular control channel was not available. If a subject learned that altitude could be controlled by either collective or pitch control, when both were present, he would be at a disadvantage during those conditions when only pitch control was effective. In the actual helicopter situation, lateral and horizontal translations are not obtained in a direct manner, but rather are the temporal consequence of maintaining certain flight attitudes, involving pitch, roll and altitude control. Being complex and attainable in a variety of control combinations, they would produce an unnecessary ambiguity to the test situation. For the purposes of the present study, an invariant relationship was required between the control parameters and the display characteristics with which they were paired. Thus, the situation obtained was such that pitch control was exercised only by the forward-backward movement of the cyclic stick. Roll was introduced by a lateral movement of the same control lever. Yaw was served by the rudder pedals and altitude change could be affected only by the collective handle. By segregating the control-display channels in the above manner, the diversification of control patterns was, in consequence, avoided.

The simulator cabin is shown in Figure 1. The windows and windscreens of the cabin were opaque. Cabin lighting was provided of an intensity sufficient for ease of control, but was not of a brightness to interfere with reading the Television monitor display. During testing, the subjects wore a helmet containing microphone and headphones. This allowed them to communicate with the experimenter who sat at a console some distance from the simulator cabin. During testing, a noise resembling that made by the rotor and engine of a helicopter was heard in subject's earphones. This, along with the opacity of the windows, was sufficient to isolate the subject from extraneous stimulus cues, and insured their reliance upon the display and upon those proprioceptive cues provided by the movement of the platform.

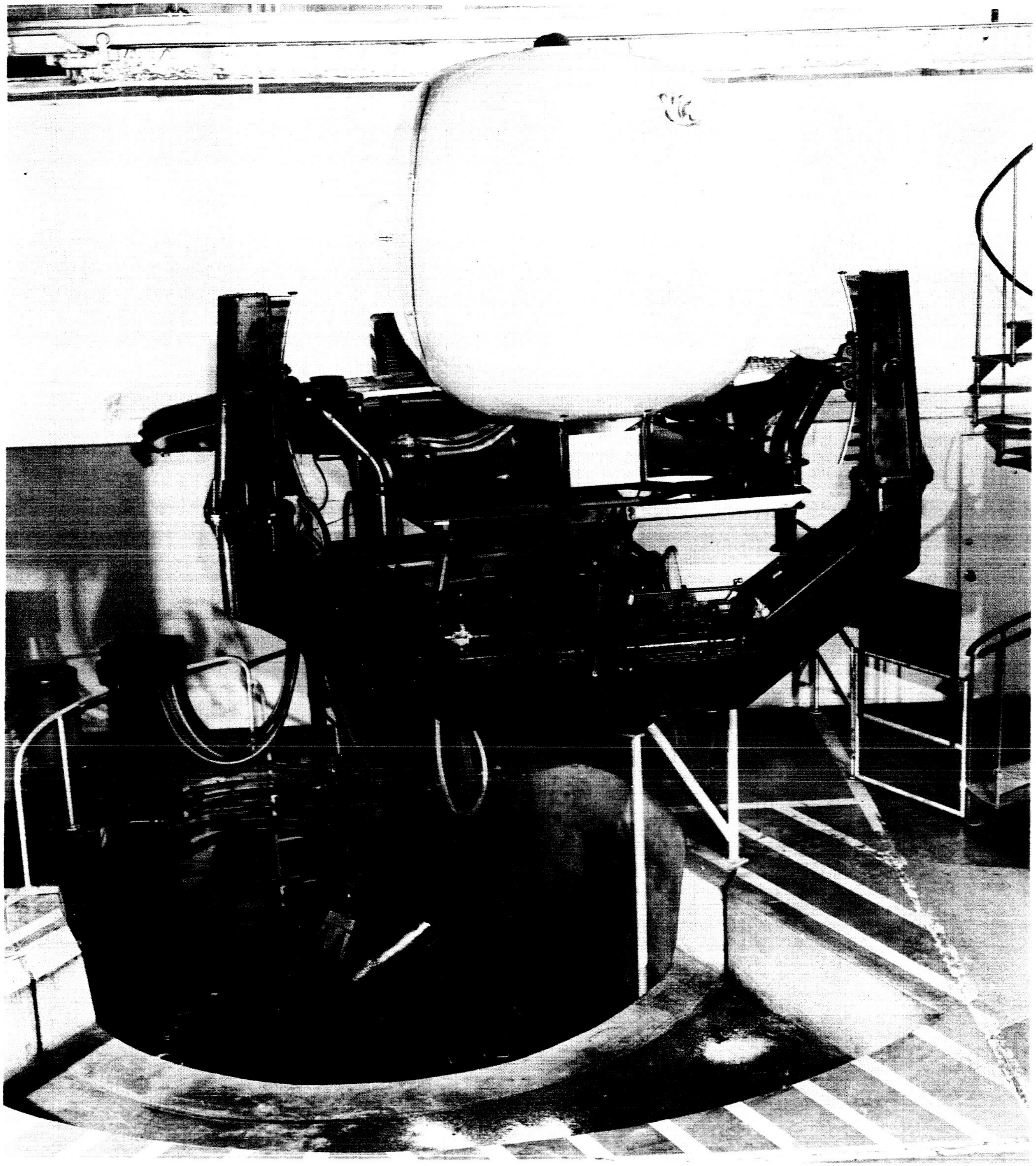

FIGURE 1. SIMULATOR MOUNTED ON DYNAMIC PLATFORM

Display: The contact analog vertical display was used by the subjects for the control of both attitude and altitude. The display image is shown in a characteristic position in Figure 2. The display signal generator system provided a variety of inputs in video form to the cockpit display. Computed attitude information in pitch and roll were displayed in the form of an earth stabilized horizon. The transformation of earth coordinate positions in the appropriate display screen coordinates were computed, utilizing signals integrated from control and noise inputs. The horizon which appeared as a white line on the upper margin of the grid could be used by the subject to ascertain pitch and roll attitude. The impression was much as though the operator was looking forward through the wind screen of a hovering helicopter. If the horizon line moved up the scope, this was correctly judged as denoting a nose down condition. Down scope indicated a nose up situation. A clock-wise rotation of the horizon line was appropriately responded to as a left roll. A rotary movement in the opposite direction was indication of a right roll.

Clouds appeared above the horizon. These had a three-fold purpose. First, they contributed to the pictorialization of the display. Second, they could be used for pitch and roll control in those conditions when the aircraft was pitched up so that the horizon line was not in the picture. Finally, the clouds were of use in yaw control since a rotary motion about this axis would cause them to shift across the screen.

Below the horizon line, a grid (intersecting lines perceptually orthogonal from an oblique visual reference converging at infinity) was shown. The size of the squares varied with the altitude as did the figure-ground relationship. From 0 to 96 feet the grid was composed of white lines on black background. As the altitude increased within this interval, the size of the squares decreased. At 96 feet the squares recovered their initial size but were then made up of white lines on a dark background. With increasing altitude, the squares again decreased in size until a heighth of 768 feet was obtained. At this point, the figure-ground relationship was again reversed and the squares recovered their full size only to diminish as still more altitude was obtained. Thus, altitude was displayed through three magnitude segments by changing the figure-ground arrangement. Within each segment, the size of the squares varied, becoming smaller as the upper limit of the segment was approached.

The grid was also used to display yaw. The perceptually paralleled lines converged at the four cardinal points of the compass. A change in yaw would produce a rotative movement of the grid which was seen as pivoting about the stabilized earth position.

Computer: The computer facility is shown in Figure 3. It was a Berkeley EASE model 1,000 analog computer. It provided the control-display dynamic link. Its outputs were also used to drive the dynamic platform upon which the simulator cabin was mounted. The experimenter sat at the console shown in the picture. He was

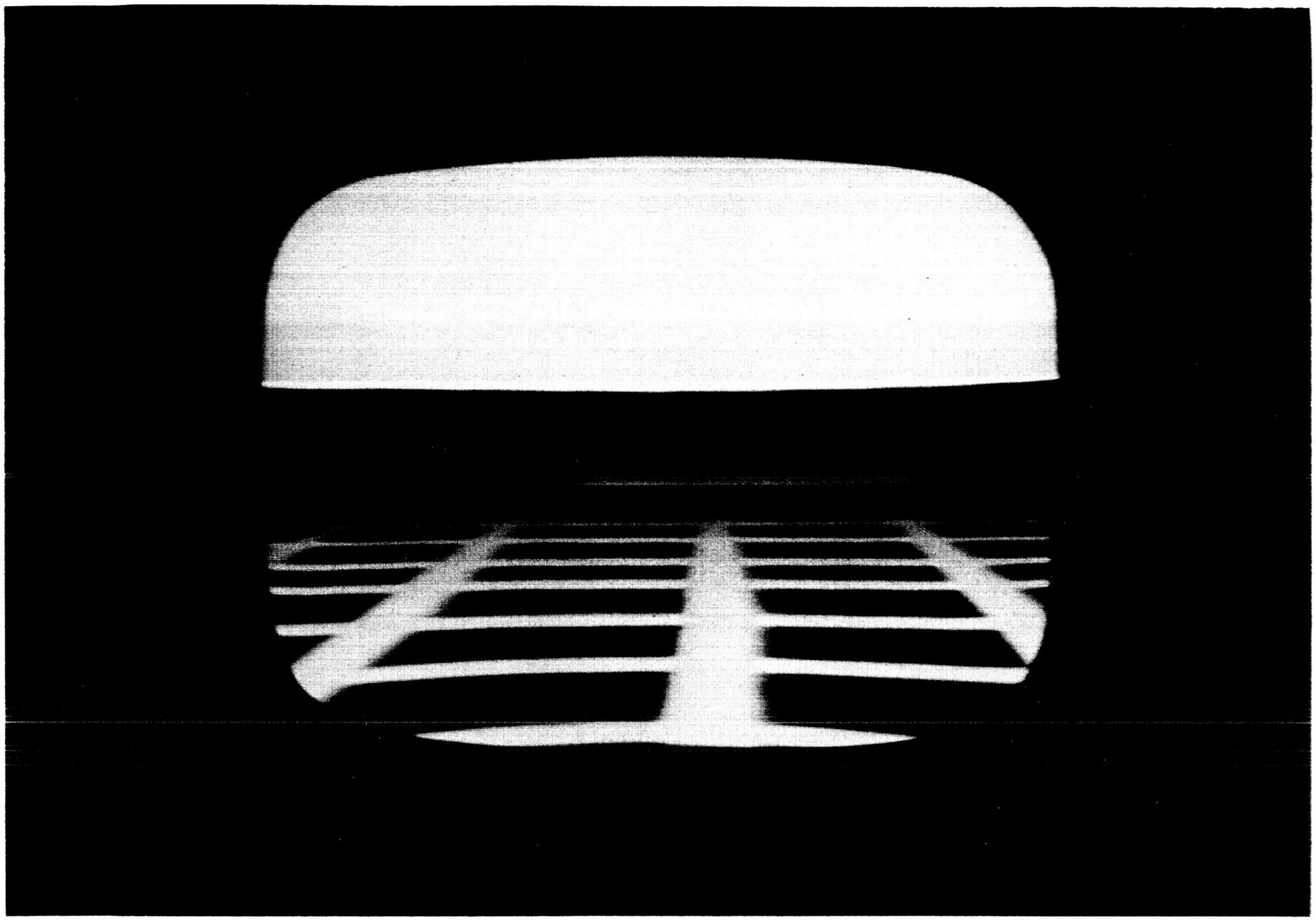

FIGURE 2. ILLUSTRATION OF CONTACT ANALOG DISPLAY

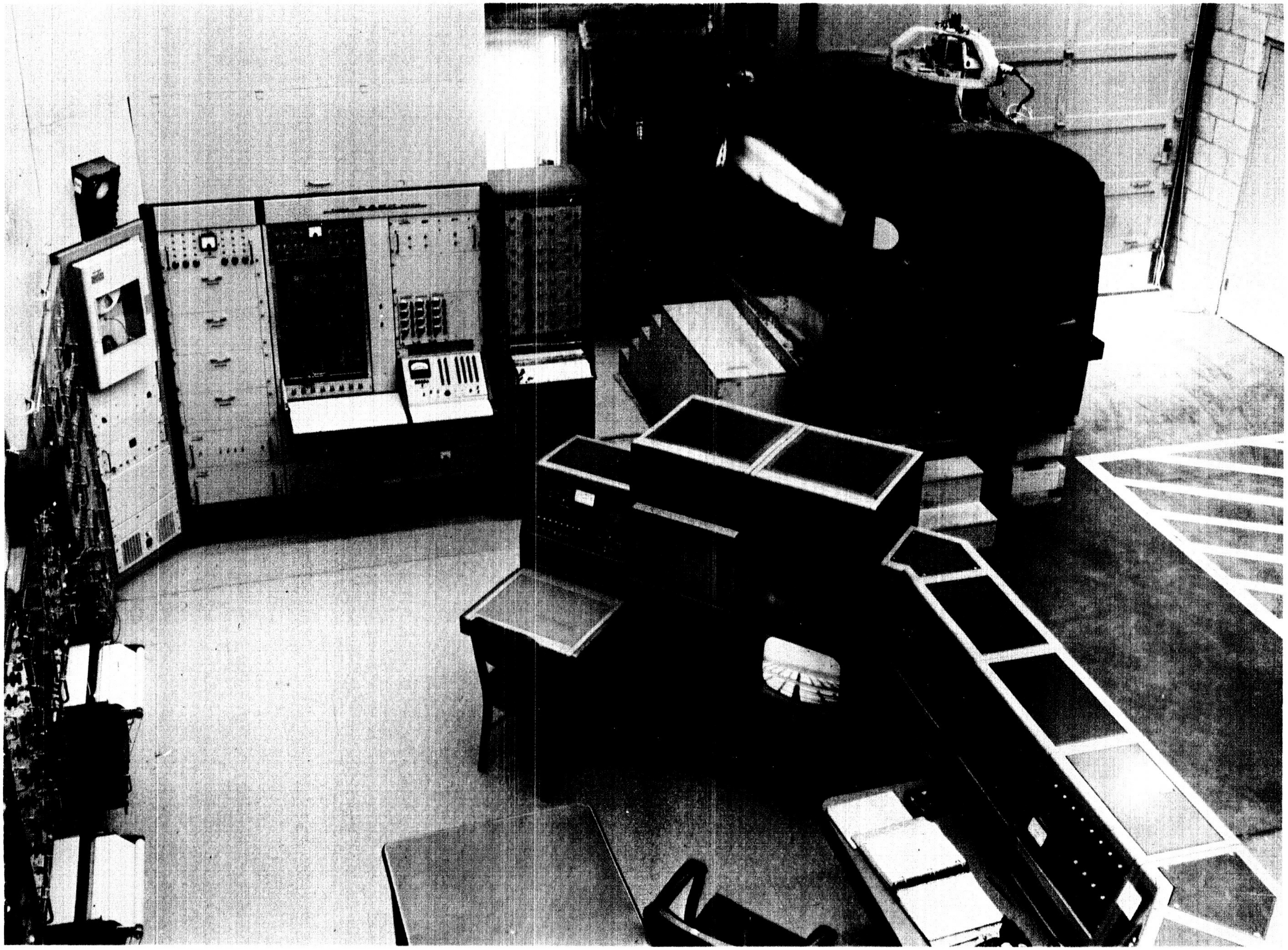

FIGURE 3. EXPERIMENTER'S CONSOLE WITH A PORTION OF THE COMPUTING AND RECORDING FACILITIES SHOWN

provided with a TV monitor which registered the same picture that was seen by the subject. In addition, the operator had an intercom system for communication with the subject and the necessary equipment for the control of the experimental procedures.

Recording: The data were recorded on magnetic tape. Depending on the experimental condition, registration was made on as many as nine channels at a time. Error measurement consisted of the following: momentary error in (1) pitch, (2) roll, (3) yaw, and (4) altitude. Activity measurements were also determined; these were acquired for (5) fore and aft position of cyclic stick, (6) lateral position of the same control handle, (7) rudder pedal position and (8) collective stick position. In addition, (9) a voice channel from the intercom was obtained which identified the condition and number of each trial.

Except for the voice channel, these records were subsequently converted to digital form before analysis was undertaken. The analysis is described in the following section.

Procedure

There were fifteen conditions of testing used in the experiment. These differed only in the combination of tasks involved. All combinations of the four tracking tasks taken one or more at a time were included. This is shown in Table I. The list of tasks for the fifteen conditions of the experiment - roll, pitch, yaw and altitude - are symbolized as R, P, Y and A respectively.

Considered from a different point of view, there were four levels of task difficulty. On the most simple level, each task was presented alone. On the second level, each task was presented in the presence of each remaining task. As a third level, each task was presented in concert with two other tasks. Finally, there was a fourth level in which all tasks were given simultaneously. If the assumptions could have been made that the tasks were of equal difficulty (as determined by the amount of time the subject spent in control of each) and that the tasks were neutral to one another with reference to operational interference, then the data could be examined using the classifications of (1) simple task, (2) task with one additional task, (3) task with two additional tasks, etc. It is unlikely that the first assumption can be made and it is certain that the second cannot. It is, in consequence, necessary that each combination be analyzed separately.

It may be correctly judged that in producing individual analysis, the initial aim of the study (description of performance as a function of task loading) is thereby sacrificed. While this may be unfortunate, little is to be gained by the pretense that task loading is in some simple relationship with the number of tasks given.

TABLE I

Condition	Tasks	Measurements Taken: Error	Measurements Taken: Control Position
1	R	Momentary error in roll	Lateral Cyclic
2	P	Momentary error in pitch	Fore-Aft Cyclic
3	Y	Momentary error in yaw	Foot Pedal
4	A	Momentary error in altitude	Collective
5	R+P	Momentary error in roll Momentary error in pitch	Lateral Cyclic Fore-Aft Cyclic
6	R+Y	Momentary error in roll Momentary error in yaw	Lateral Cyclic Foot Pedal
7	R+A	Momentary error in roll Momentary error in altitude	Lateral Cyclic Collective
8	P+Y	Momentary error in pitch Momentary error in yaw	Fore-Aft Cyclic Foot Pedal
9	P+A	Momentary error in pitch Momentary error in altitude	Fore-Aft Cyclic Collective
10	Y+A	Momentary error in yaw Momentary error in altitude	Foot Pedal Collective
11	R+P+Y	Momentary error in roll Momentary error in pitch Momentary error in yaw	Lateral Cyclic Fore-Aft Cyclic Foot Pedal
12	R+P+A	Momentary error in roll Momentary error in pitch Momentary error in altitude	Lateral Cyclic Fore-Aft Cyclic Collective
13	R+Y+A	Momentary error in roll Momentary error in yaw Momentary error in altitude	Lateral Cyclic Foot Pedal Collective
14	P+Y+A	Momentary error in pitch Momentary error in yaw Momentary error in altitude	Fore-Aft Cyclic Foot Pedal Collective
15	R+P+Y+A	Momentary error in roll Momentary error in pitch Momentary error in yaw Momentary error in altitude	Lateral Cyclic Fore-Aft Cyclic Foot Pedal Collective

It can be seen that under Condition 1 that only roll was present. In other words, the display and the platform remained stable and at the null position for pitch, yaw and altitude since only roll inputs were present and only the lateral movement of the cyclic was effective. During this condition, records were obtained only of roll error and of the lateral movement of the cyclic. For Condition 8, both pitch and yaw varied with the subjects' inputs and the forcing functions on these channels. Similarly, records were obtained only of pitch and yaw error and of the fore-aft movement of the stick along with the movement of the foot pedals. Only under Condition 15 were all channels active, each requiring the attention of the subject. When this condition was present, all eight quantitative magnetic tape records were made.

On each day of testing, each subject received each of the 15 conditions described above. Each condition was given for a two minute interval. The non-pilots were given 15 days of testing while the pilots received only one day. The order of presentation of conditions varied between subjects and between days as shown in the latin square of Table II. On any given day of testing, each subject received an order of condition presentation that was different from that received by any other subject. During the 15 days, however, each subject received each sequence of conditions. Subject No. 2, for example, received order 15 on the first day, order 5 on the second, etc. The orders were made up of different sequential arrangements of the 15 conditions shown in Table I. In Table III, the daily sequence of conditions for the 15 orders is shown for Subject No. 2. On the first trial of the first day, the subject was required to control for roll, pitch, yaw and altitude simultaneously. On the second trial, only altitude control was required, etc. On the second day, his first trial consisted of controlling for pitch, yaw and altitude. The above procedure was followed to cancel the effects of order for each day of testing. The pilots followed the same procedure, but for the single day only.

Referring again to Table I, it will be observed in the fifteen conditions each control parameter occurred on 8 occasions. Thus, roll was obtained on Conditions 1, 5, 6, 7, 11, 12, 13 and 15. There being four such control parameters and 2 measurements of each parameter (error and control position) along with 15 subjects for 15 days, this produced 14,400 data units of 2 minutes duration for the non-pilots and 960 such scores for the pilots. Various transformations were made upon the data. These will be subsequently described.

Prior to the first days testing, a subject was given a general introduction to the apparatus and told what the procedures would be. He was told, for example, that the pedals controlled for yaw, the collective for altitude, etc. He was instructed to fly straight and level on a North heading at 50 feet of altitude. He was then given 2 minutes experience controlling each channel separately. These pre-trial procedures were used for both pilots and non-pilots; except in the case of the former group, the discussion of the relationship between control movement to display movement was considerably simplified.

SUBJECT

DAY	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	8	15	2	12	5	14	7	3	11	1	9	10	13	6	4
2	9	5	11	7	6	4	8	10	3	14	12	1	2	13	15
3	4	12	8	3	11	15	5	9	7	13	10	14	1	2	6
4	1	9	7	2	12	3	14	13	10	6	5	8	15	4	11
5	14	8	15	10	3	2	9	7	1	11	4	6	12	5	13
6	3	10	5	8	15	1	6	11	14	2	13	12	4	7	9
7	11	1	4	6	9	12	3	14	2	10	7	13	5	15	8
8	5	14	3	13	2	8	10	6	4	7	15	9	11	1	12
9	12	3	10	11	7	13	4	15	6	9	2	5	14	8	1
10	13	6	12	14	8	11	1	2	15	5	3	4	7	9	10
11	6	11	1	15	10	5	12	4	13	3	8	2	9	14	7
12	2	13	9	5	4	6	11	12	8	15	1	7	10	3	14
13	7	2	13	1	14	10	15	5	9	4	6	11	8	12	3
14	10	7	6	4	1	9	13	8	5	12	14	15	3	11	2
15	15	4	14	9	13	7	2	1	12	8	3	5	6	10	5

TABLE II - Latin Square Showing The Sequence Of Orders Of Presentation Given Each Of The Non-Pilot Subjects

DAY	ORDER	TRIAL														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
7	1	PY	RPYA	P	RPA	RP	PYA	RA	Y	RPY	PA	RYA	R	YA	A	RY
13	2	PA	RP	RPY	R	RY	A	PY	YA	Y	RYA	RPA	RA	PYA	P	RPYA
9	3	A	RPA	PY	Y	RPY	RPYA	RP	PA	RY	PYA	RA	YA	RYA	R	P
15	4	R	PA	RA	RPYA	RPA	Y	PYA	RYA	YA	P	RY	RPY	A	PY	RP
2	5	PYA	PY	RPYA	A	Y	P	PA	RA	RYA	RP	YA	RY	RPA	RPY	R
10	6	Y	YA	RP	PA	RPYA	R	RY	RPY	RA	RPA	P	A	PY	RYA	PYA
14	7	RPY	R	A	YA	PA	RPA	Y	PYA	RP	RA	PY	P	RPYA	RY	RYA
5	8	RP	PYA	Y	RYA	P	PY	YA	RY	RPYA	R	PA	RPA	RPY	RA	A
4	9	RPA	Y	YA	RPY	RA	RYA	A	RPYA	P	RY	PYA	PY	R	RP	PA
6	10	RYA	RY	RPA	RP	PY	RPY	R	P	PYA	YA	A	RPYA	RA	PA	Y
11	11	RY	RPY	R	PYA	YA	RP	RPA	A	PA	PY	RPYA	RYA	P	Y	RA
3	12	P	RYA	PA	RA	A	RY	RPY	RPA	PY	RPYA	R	Y	RP	PYA	YA
12	13	RA	P	RYA	PY	PYA	YA	RPYA	RP	R	A	Y	PA	RY	RPA	RPY
8	14	YA	RA	RY	P	R	PA	RYA	PY	A	RPY	RP	PYA	Y	RPYA	RPA
1	15	RPYA	A	PYA	RY	RYA	RA	P	R	RPA	Y	RPY	RP	PA	YA	PY

TABLE III - Sequence Of Conditions For Each Order Given To Subject No. 2

V. RESULTS

As previously related, two measures were continually enregistered for each of the four control parameters (roll, pitch, yaw and altitude). These consisted of momentary error (deviations from a level attitude on a north heading at 50 feet) and control position. These data were converted into digital form by sampling the magnetic tapes at the rate of 80 times a second. These numeric values were then used in the construction of the following measures.

$$(1) \qquad e = \int (a - d)dt$$

This value represents the total algebraic error. Since the momentary measurements were about as often negative as positive, these scores were close to zero unless a systematic error persisted. The only value to this statistic was in the production of the standard deviation (number 5 below).

$$(2) \qquad w = \int \frac{d\ (s - d)}{dt}\ dt$$

This measure was integrated velocity of the control lever. In the analysis it was used as an index of the subject's activity or work input.

$$(3) \qquad /e/ = \int /a - d/dt$$

This was the absolute (accumulated without reference to sign) integrated error.

$$(4) \qquad e^2 = \int (a - d)^2\ dt$$

In this measure, the momentary errors were squared before being integrated. This, of course, produces a relatively high penalty for large errors.

$$(5) \qquad \sigma = \sqrt{\frac{e^2}{N} - \frac{(e)^2}{N}}$$

This also tends to assess a disproportionate penalty for large errors, but unlike /e/ and e^2 it removes the penalty for constant errors. It is, in consequence, closer to the subject's perception of error than are the other measurements which are based upon the experimenters criteria for error.

(6) /e/ w

This was an index of inefficiency based upon the absolute error scores. It was simply a product of measures (2) and (3) as listed above.

(7) e^2w

This was an index of inefficiency based upon the squared error. It was the product of measures (2) and (4).

(8) σ w

This index of inefficiency was obtained from the product of measures (2) and (5).

The logic for these manipulations has been touched upon in a previous section. Basically, the absolute error is obtained as being typical of the measurements that are often presented in tracking studies. The squared error is grossly representative of an attempt to formulate a more realistic scale than the linear function which is given by the absolute error. The standard deviation has the merit of the squared error technique, but it tends to consider only those errors that the subject acknowledges as would be the case if the subject consistently made a 1,000 foot error in altitude. The indices of inefficiency are presented in an effort to compare these parameters with their complimentary error scores to determine whether they produce a more sensitive test of the results.

In a gross way, the measurement σ w considers only the performance based upon that error the subject acknowledges to exist and works to eliminate. The measurement σ also denotes such an acknowledgment, but makes no reference to subject's activity. Differences between the relative values of σ and /e/ and of σ w and /e/w are taken to indicate the presence of error that the subjects were unaware of.

These measurements were produced in both absolute and relative forms. In the latter, the effect of subject differences was removed. These values are shown for the various treatment combinations in Figures 4 through 7 for absolute error /e/ only. The purpose of presenting these graphs is to illustrate the fact that as training progresses the ratio of task difficulty as a function of number of associated tasks approaches unity. This phenomena was also evident for the other methods of measurement, but it was not regarded as necessary that they be presented here.

In Figures 8 through 31, the difficulty of the treatment conditions is shown for the absolute values of the data. As an additional convenience, the results are broken down into blocks of five days each. Of these figures, the first six pertain to roll, the second six to pitch, etc. Each method of measurement is given a separate graph. With respect to these curves, the following observations are pertinent.

Pitch and roll show systematic improvement through the fifteen days of testing. There is evidence, however, that the slopes of the curves tend to attenuate as the experiment progresses.

Yaw shows improvement during the first ten days, but little thereafter.

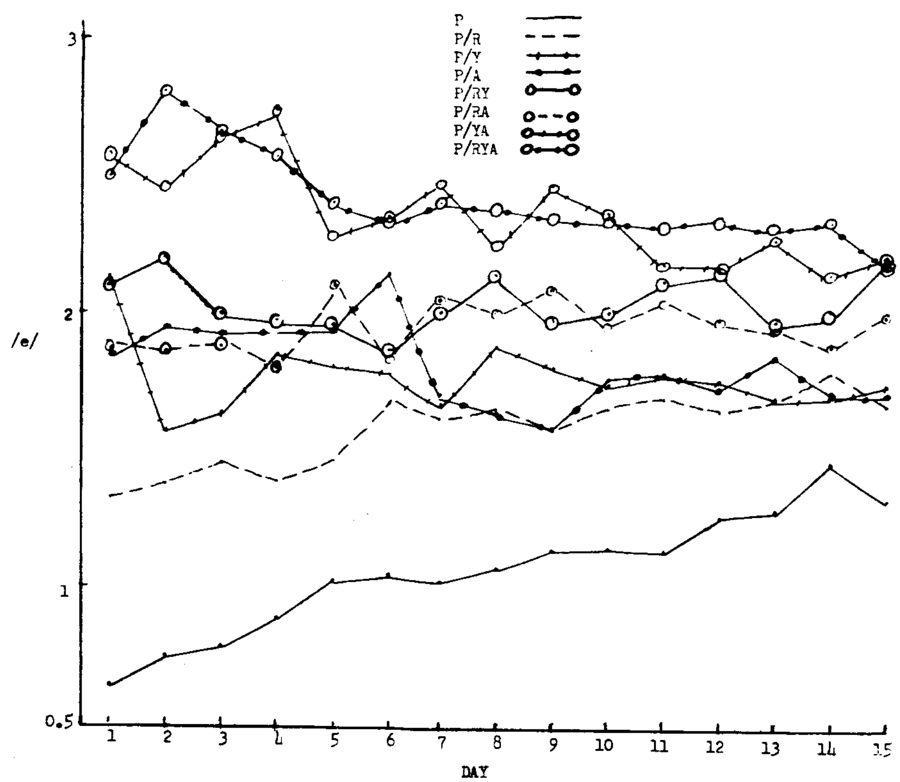

Figure 4 Absolute error (relative scores) for the eight conditions under which pitch was present for the fifteen days of testing.

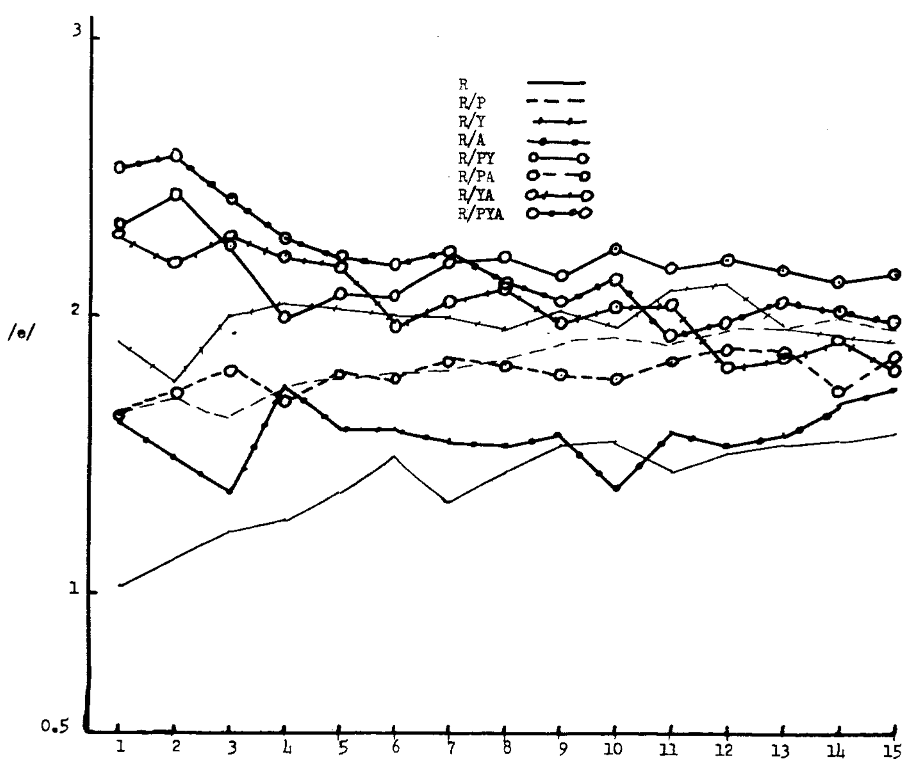

Figure 5 Absolute error (relative scores) for the eight conditions under which roll was present for the fifteen days of testing.

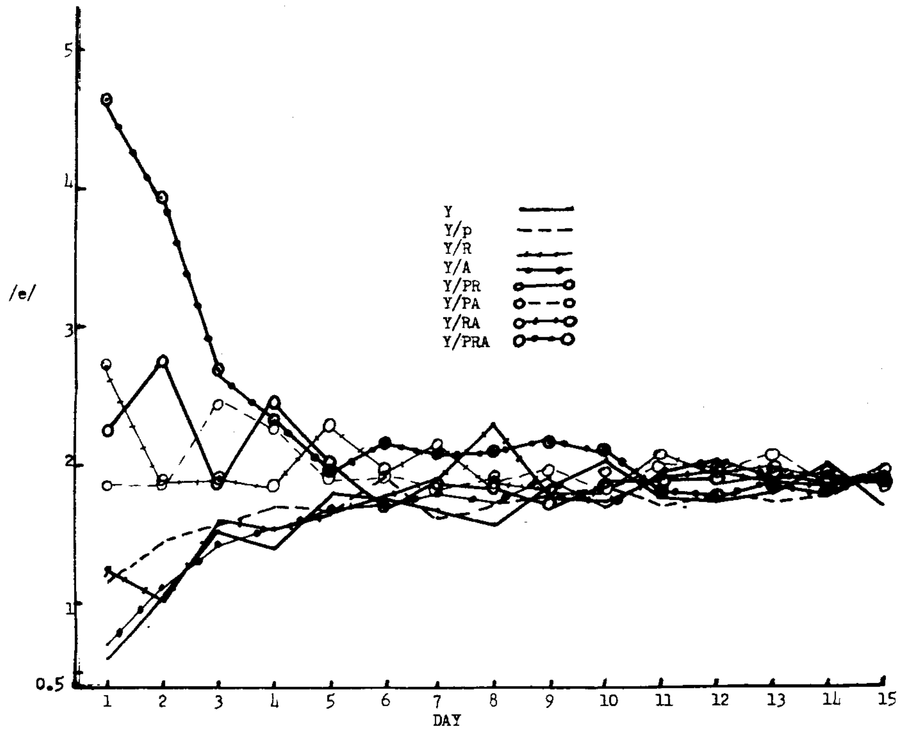

Figure 6 Absolute error (relative scores) for the eight conditions under which yaw was present for the fiteen days of testing.

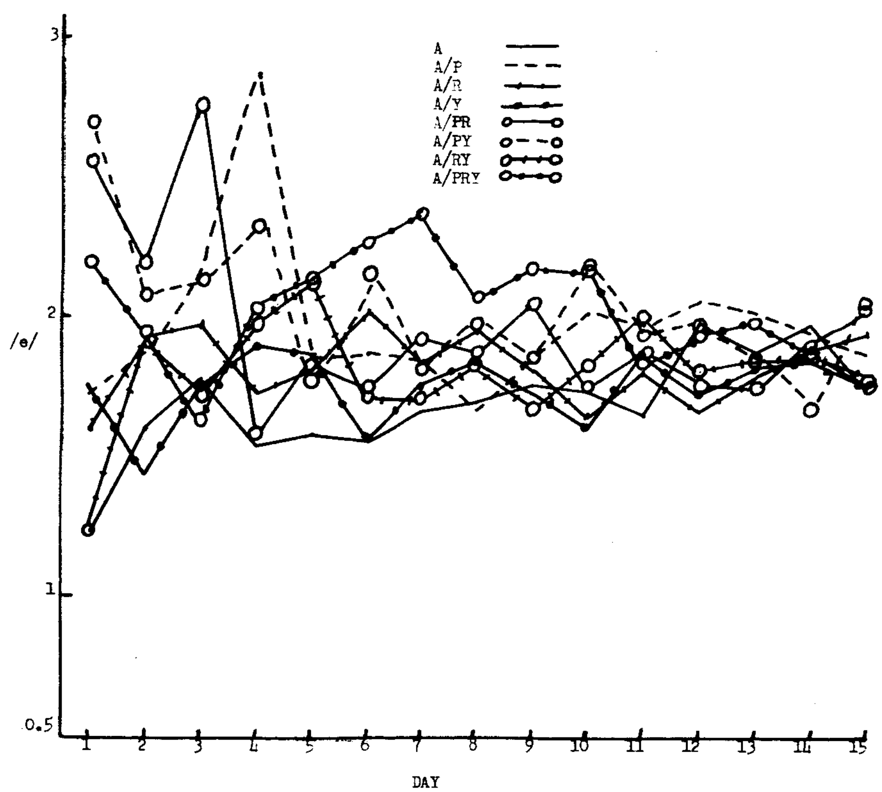

Figure 7 Absolute error (relative scores) for the eight conditions under which altitude was present for the fifteen days of testing.

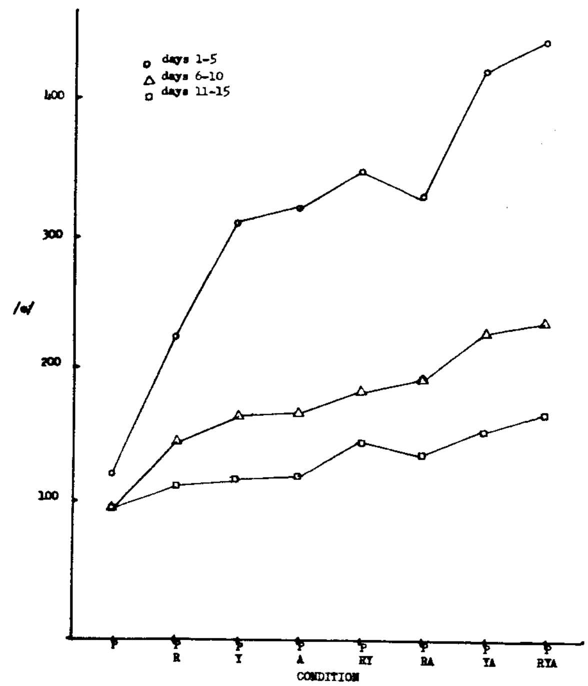

Figure 8 Absolute error /e/ for the eight conditions of pitch.

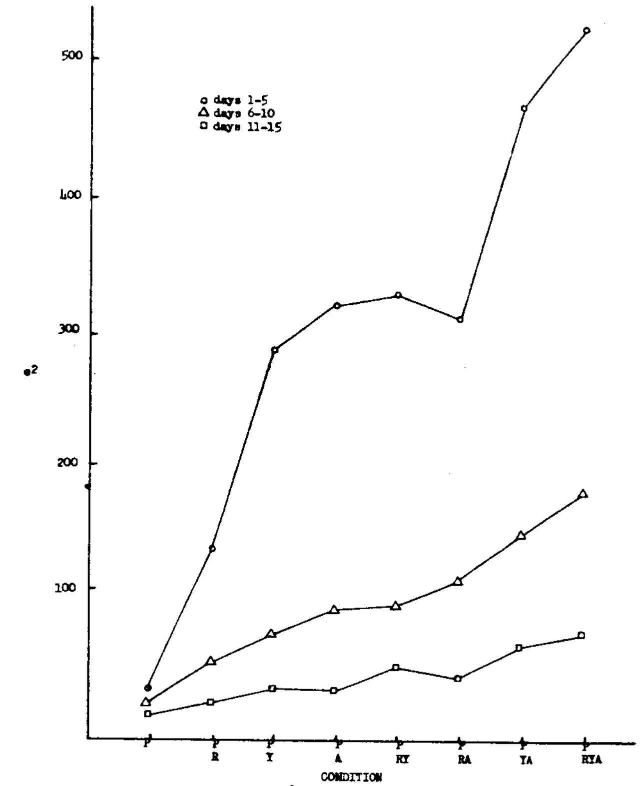

Figure 9 Squared error e^2 for the eight conditions of pitch.

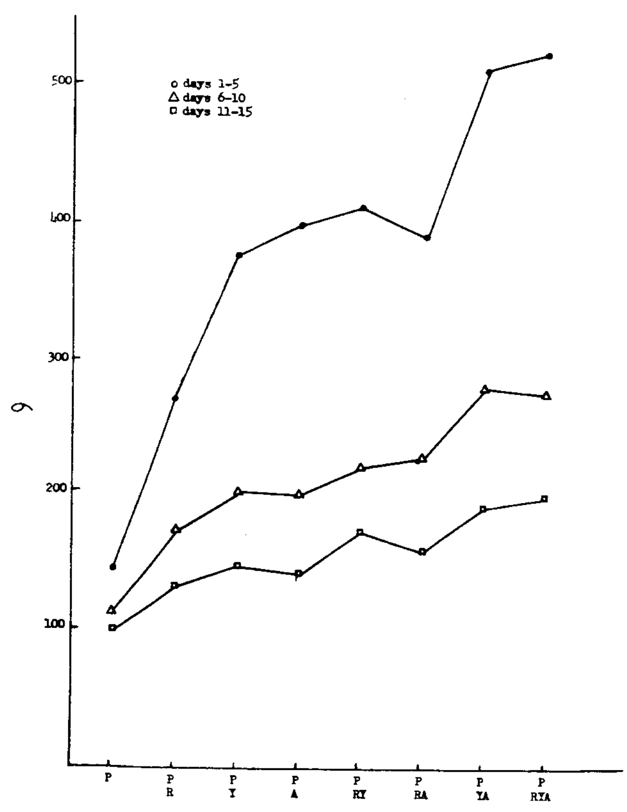

Figure 10 Standard deviation σ for the eight conditions of pitch

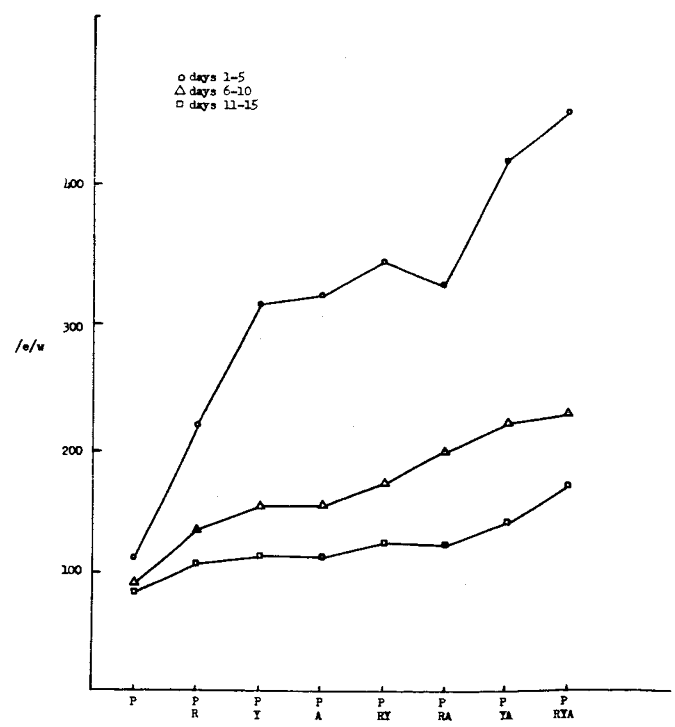

Figure 11 Inefficiency /e/w for the eight conditions of pitch.

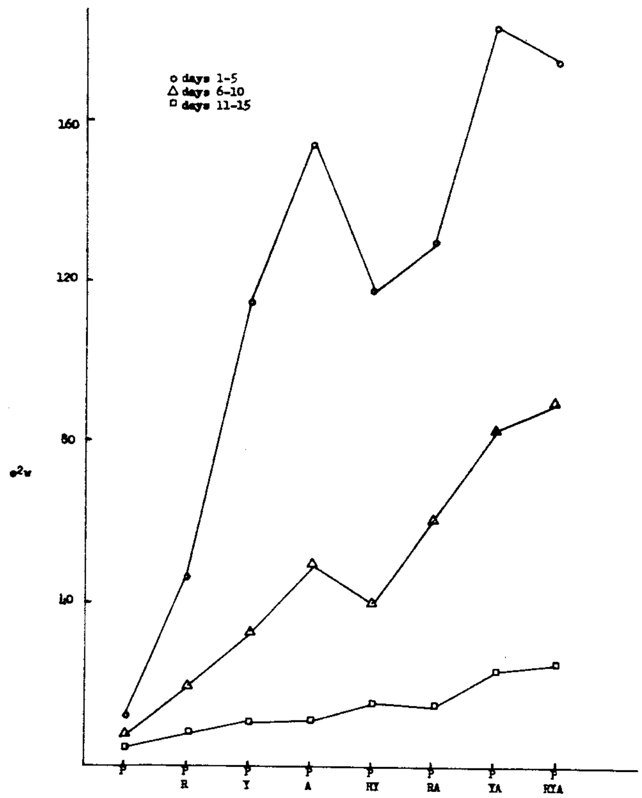

Figure 12 Inefficiency e^2w for the eight conditions of pitch.

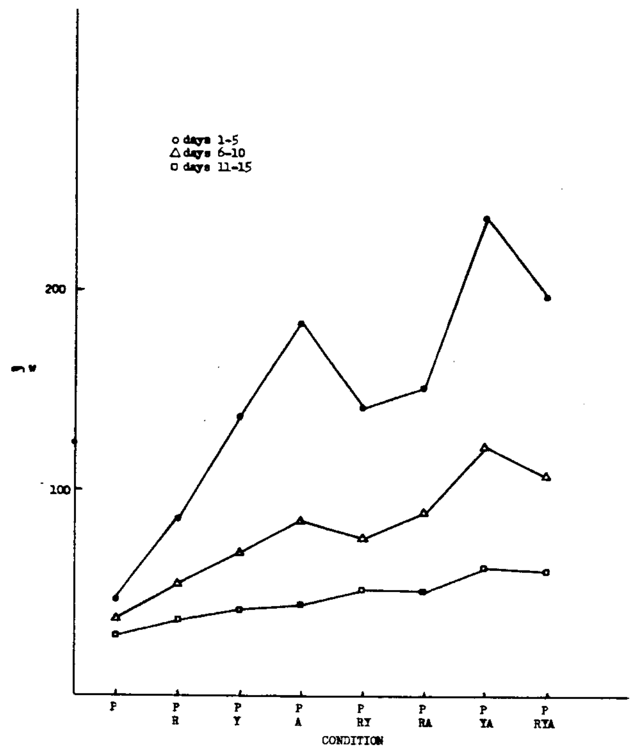

Figure 13 Inefficiency for the eight conditions of pitch.

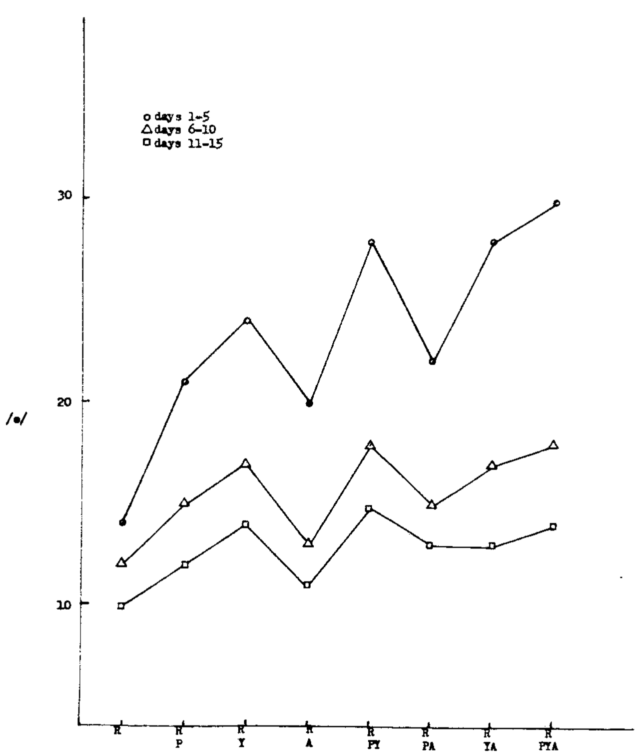

Figure 14 Absolute error /e/ for the eight conditions of roll.

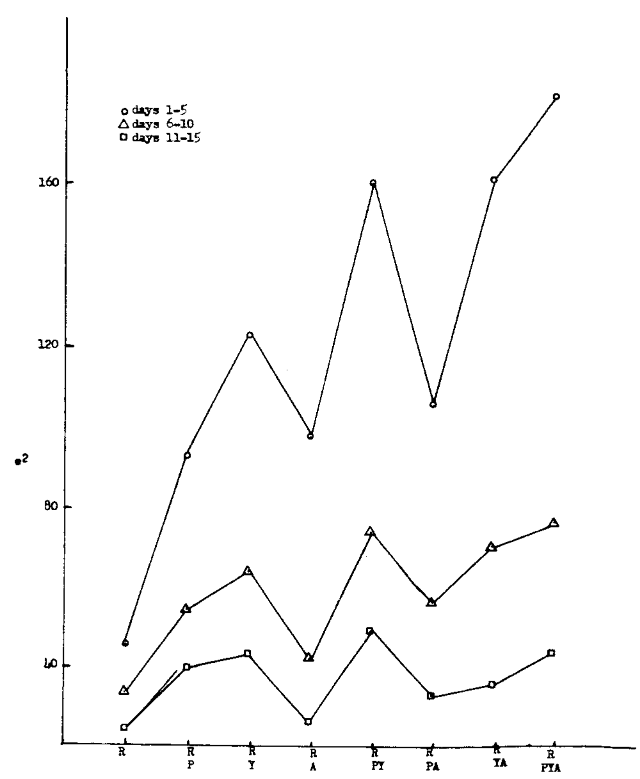

Figure 15 Squared error e^2 for the eight conditions of roll.

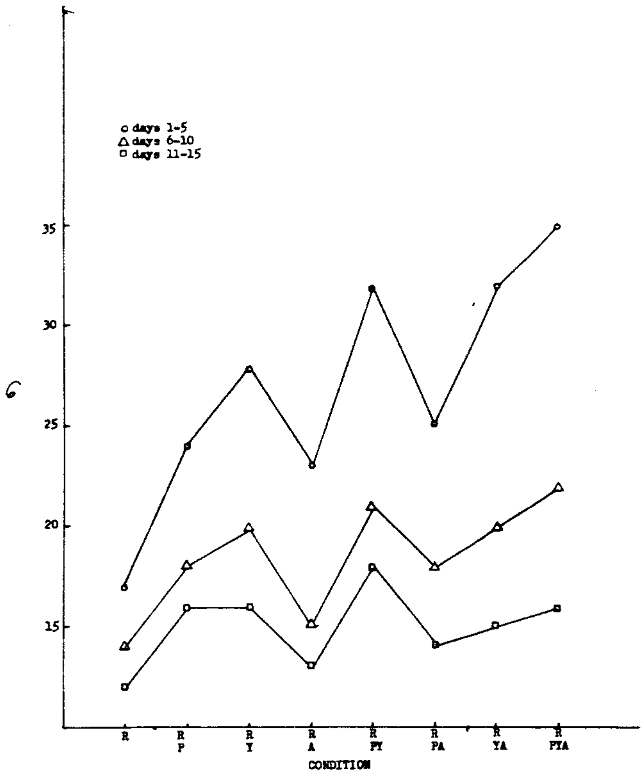

Figure 16 Standard deviation σ for the eight conditions of roll.

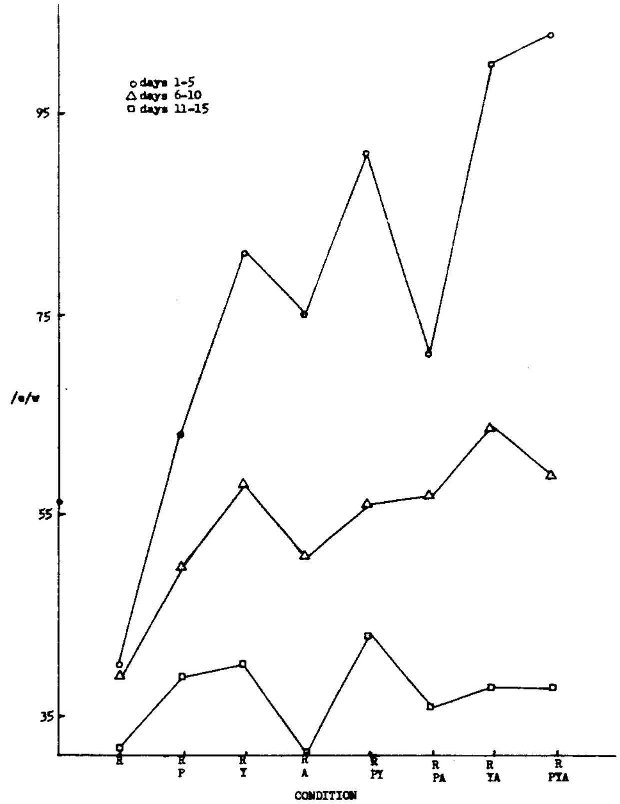

Figure 17 Inefficiency /e/w for the eight conditions of roll.

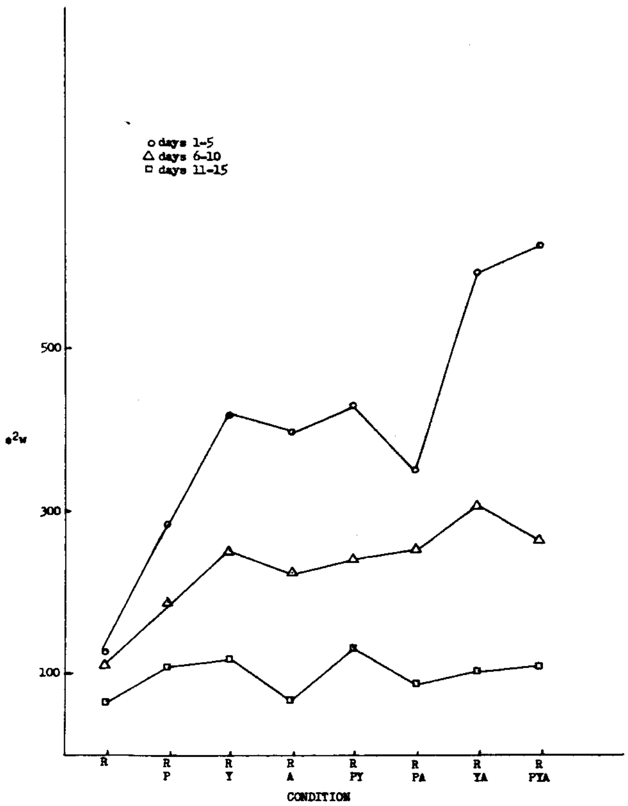

Figure 18 Inefficiency e^2w for the eight conditions of roll.

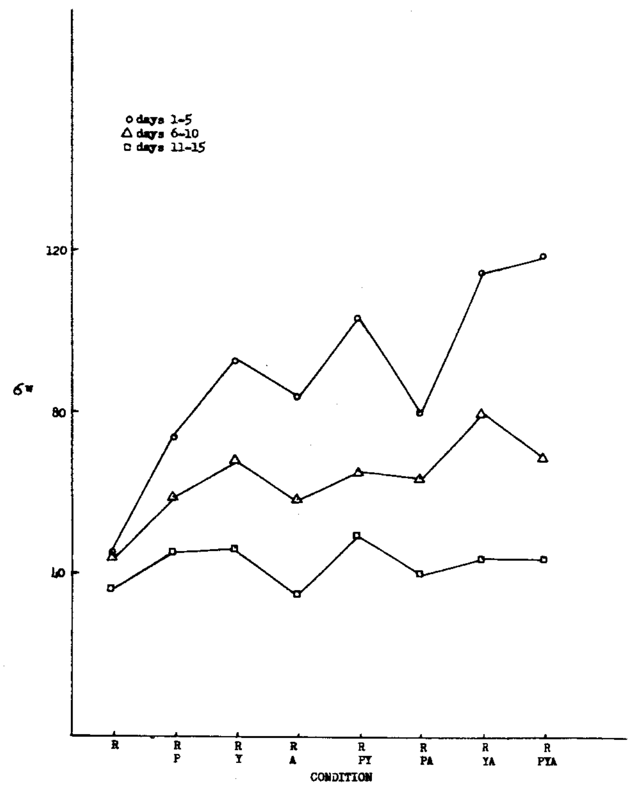

Figure 19 Inefficiency σw for the eight conditions of roll.

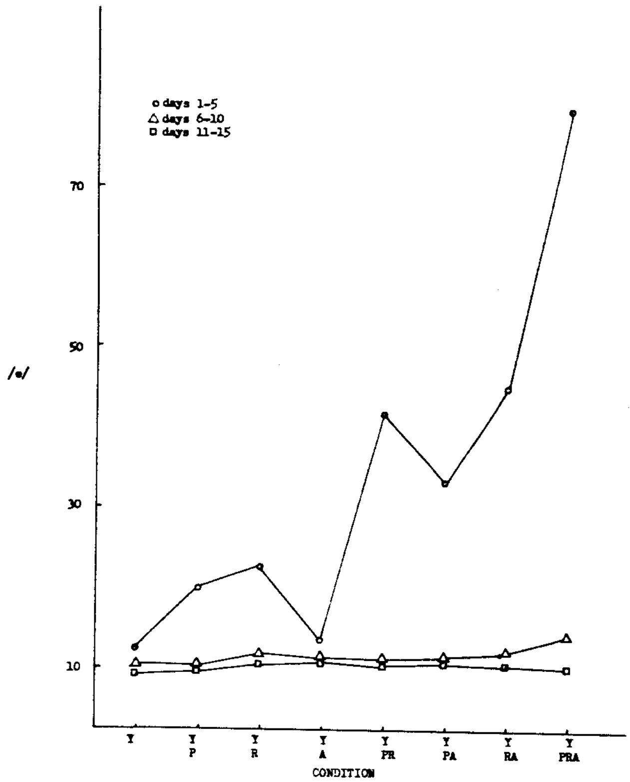

Figure 20 Absolute error /e/ for the eight conditions of yaw.

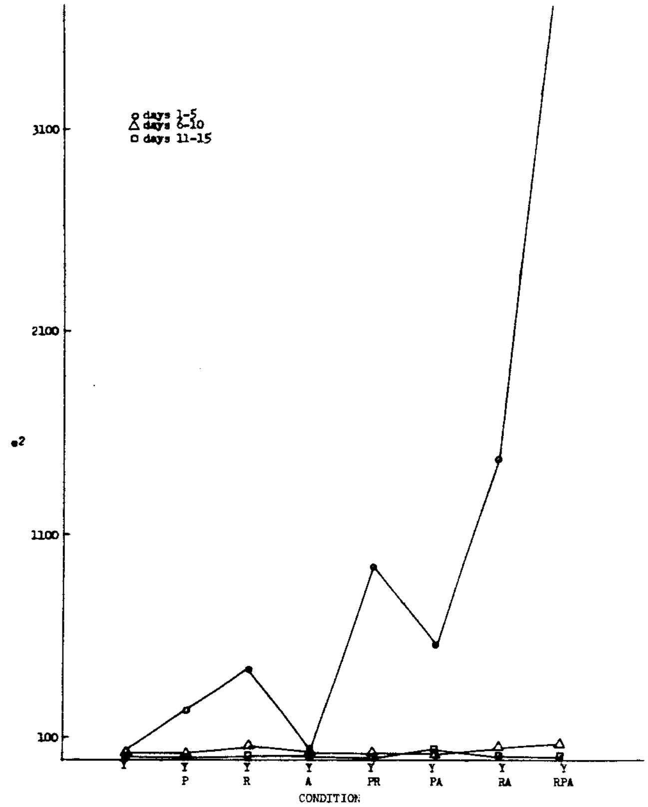

Figure 21 Squared error e^2 for the eight conditions of yaw.

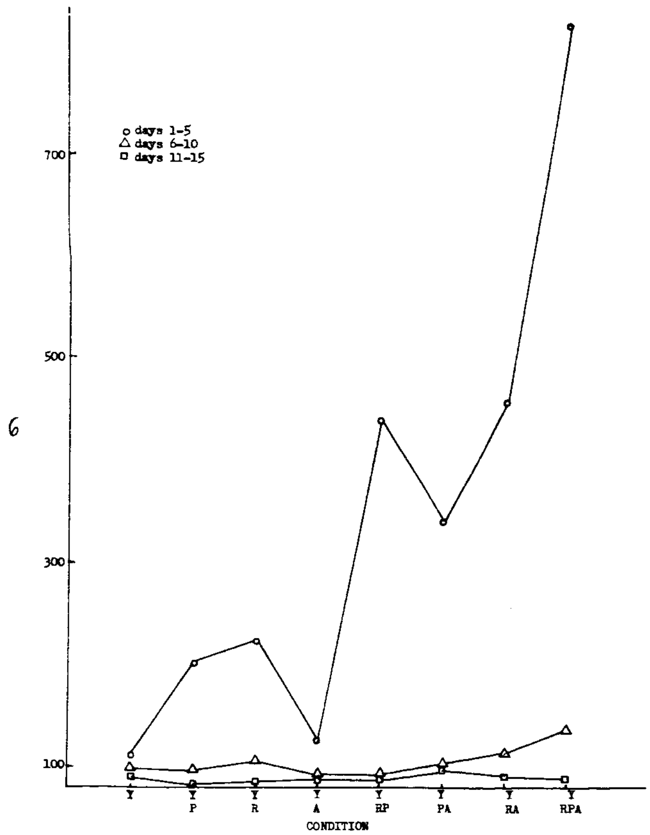

Figure 22 Standard deviation σ for the eight conditions of yaw.

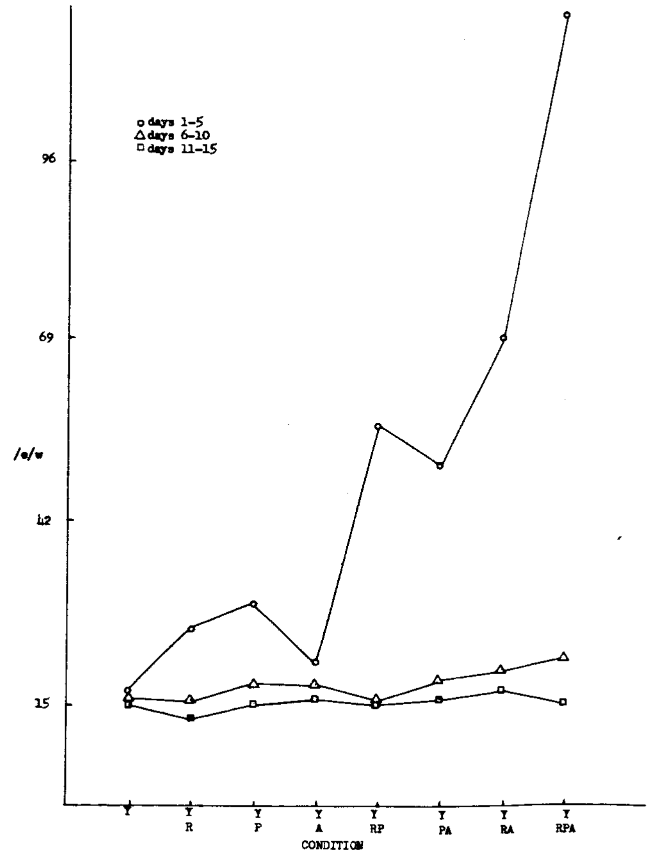

Figure 23 Inefficiency /e/w for the eight conditions of yaw.

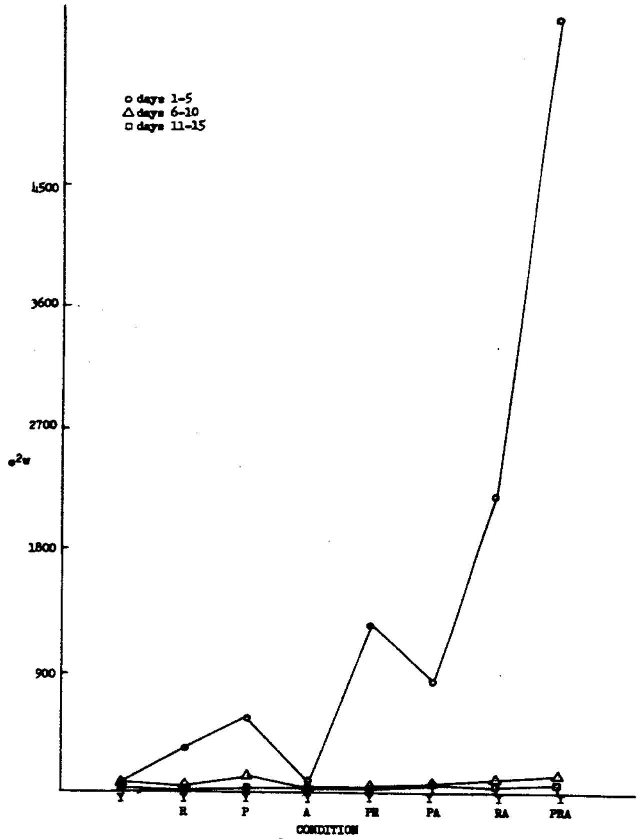

Figure 24 Inefficiency e^2w for the eight conditions of yaw.

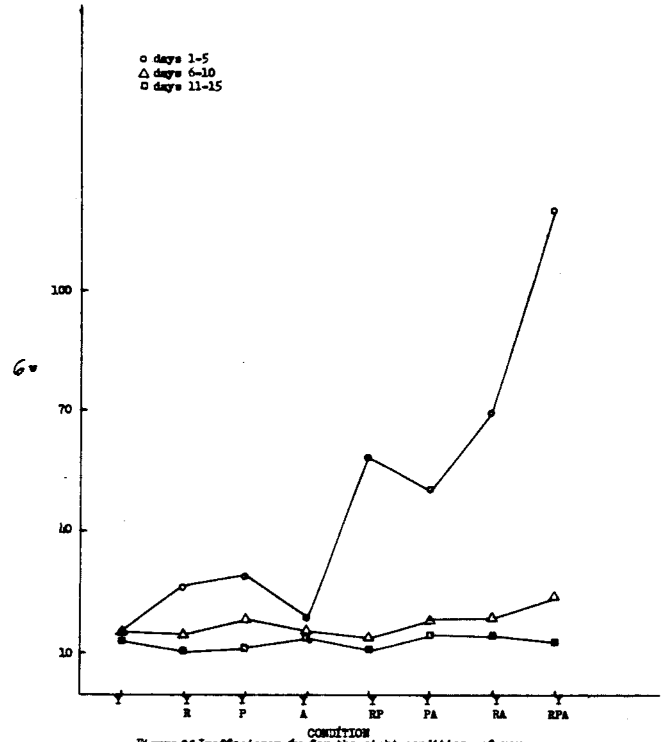

Figure 25 Ineffeciency σw for the eight conditions of yaw.

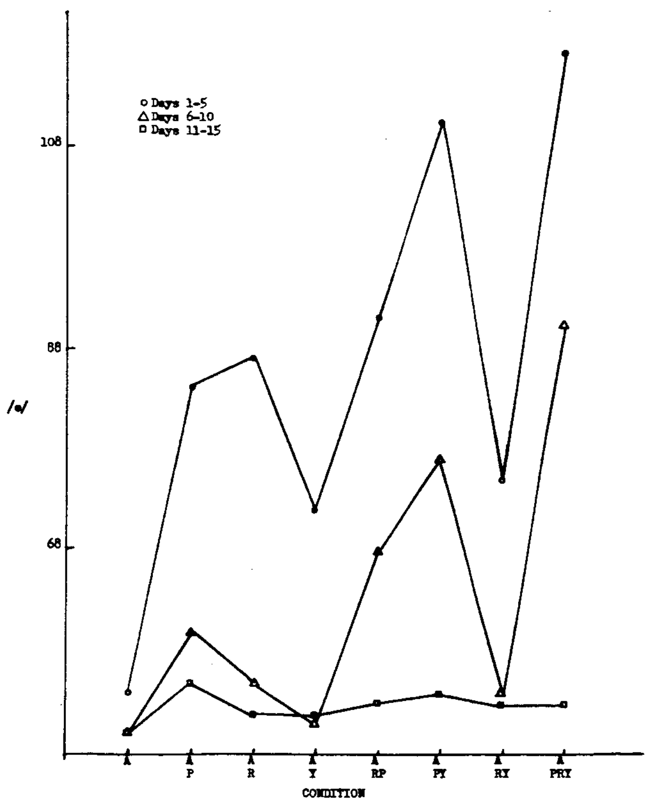

Figure 26 Absolute error /e/ for the eight conditions of altitude.

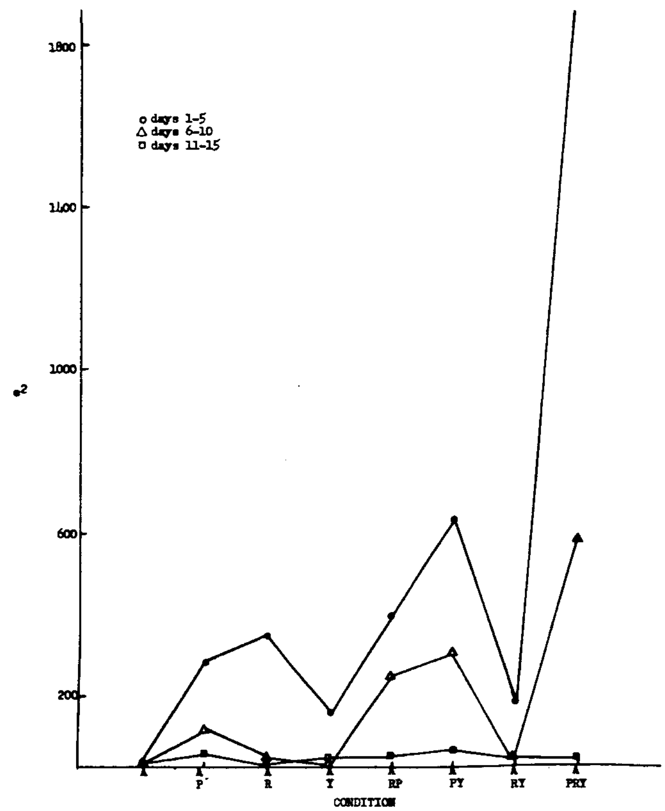

Figure 27 Squared error e^2 for the eight conditions of altitude.

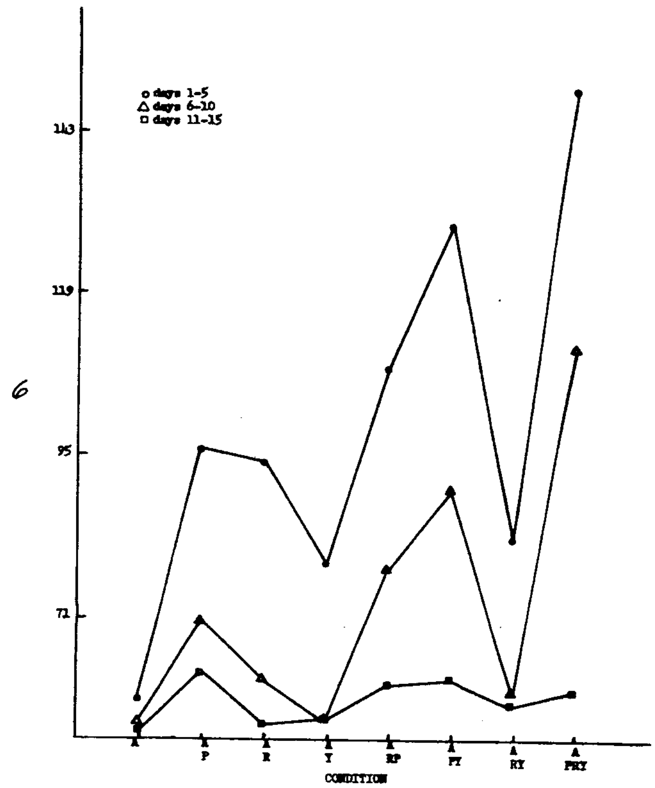

Figure 28 Standard deviation σ for the eight conditions of altitude.

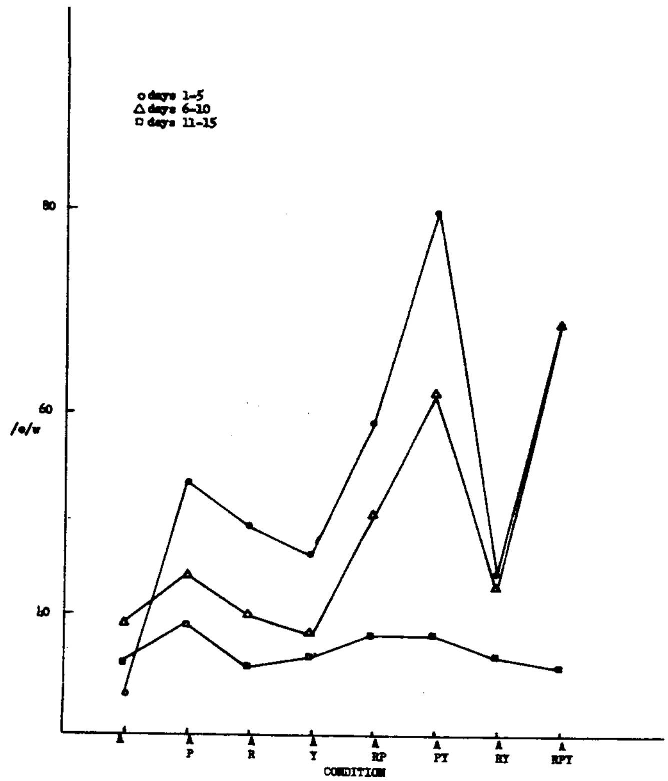

Figure 29 Inefficiency /e/w for the eight conditions of altitude.

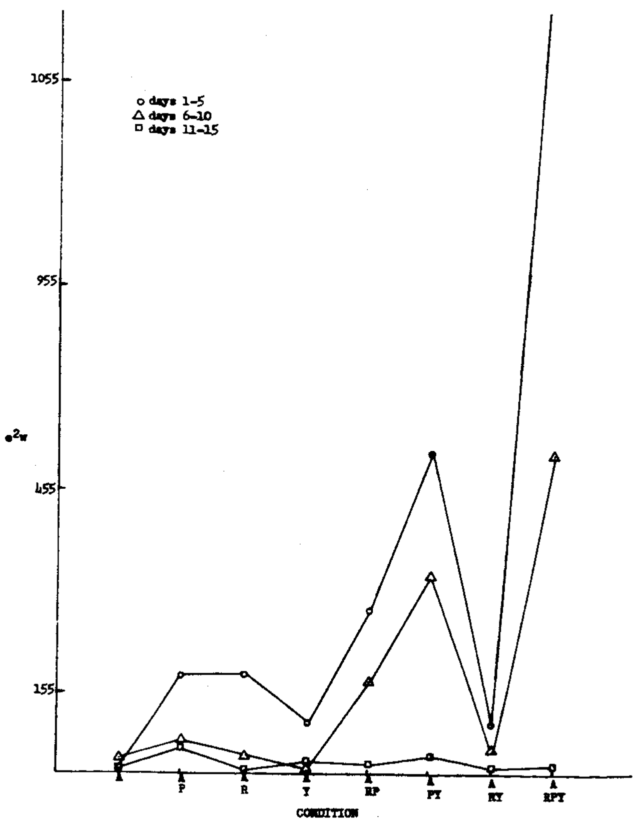

Figure 30 Inefficiency e^2w for the eight conditions of altitude.

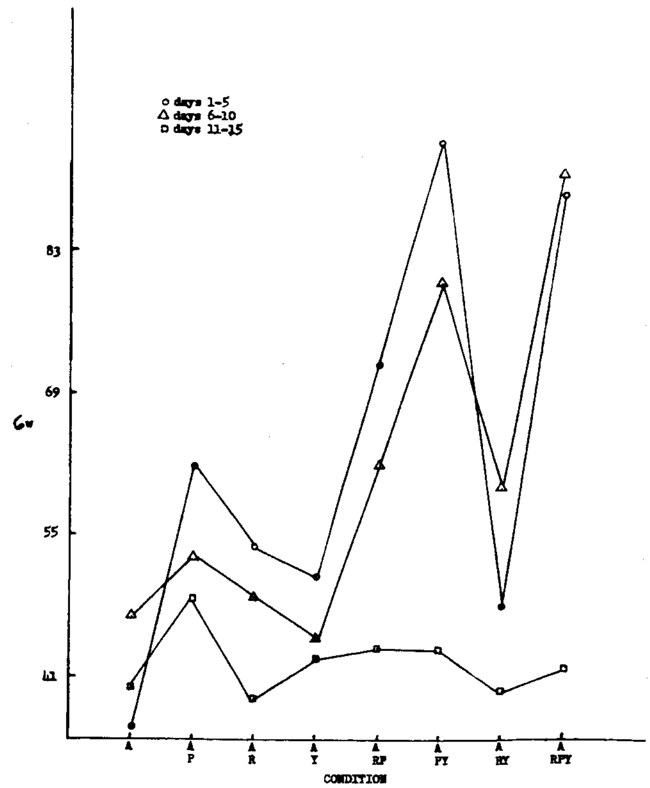

Figure 31 Inefficiency σ w for the eight conditions of altitude.

Improvement on altitude control tends to be erratic depending upon the test conditions and the type of measurement.

Generally, the measurements of inefficiency are less erratic than are the measurements of error.

As compared to the other control parameters, altitude shows large fluctuations.

The presence of altitude adversely affects pitch control. Similarly, the presence of pitch disrupts altitude control. A corresponding relationship is also to be found between roll and yaw.

In Figures 32 through 127, comparison is made of each error curve with its complimentary inefficiency curve. Since these are based upon different manipulations and since the reading of a voltage in analog recording is an arbitrary matter only the relative slopes of these curves are to be regarded as having importance.

Despite the fact that the error and inefficiency functions are quite similar, a basic difference was found to prevail. This is shown in Figures 128 through 159 where the F ratios obtained from the 48 analyses of variance (2 subject sources x 4 control parameters x 6 measurements) are illustrated. It is evident from these graphs that the variances due to treatments and days of testing are usually significant for all parameters and for all measurements. Since this was not unexpected, the results are presented in a form that will enable the reader to judge the relative significance of the differences obtained in the error measurements and the inefficiency scores.

The first eight of these graphs pertain to the pilot data while the remaining 24 relate to the non-pilot results. Since only one day of testing was accomplished on the former, the variable of day was not obtained. The F ratios for the treatment variable were obtained using the pooled mean square error derived from the subject and subject x treatment interaction as the denominator. The subject variable was tested using the subject x treatment interaction alone. In these analyses, the variables of treatments, subjects and treatments x subjects were given 7, 14 and 98 degrees of freedom respectively.

In the non-pilot data, all first order interactions as well as the subject variable were treated using the second order interaction (subject x day x condition) as the error term. These were then pooled to test the day and treatment variables. In these analyses, 1,799 degrees of freedom were exercised being divided into 14, 14 and 7 degrees of freedom for the primary effects.

The primary finding here is that σ tends to produce a higher between-conditions F ratio than does /e/ and /e/ is higher than e^2. At the same time, however, the significance of the between subject differences is reversed with e^2 being the highest.

Similarly, the measures of inefficiency produced a greater significance to the difference between conditions than did the error scores. Again, this was made possible by a less significant difference in the between subject comparisons.

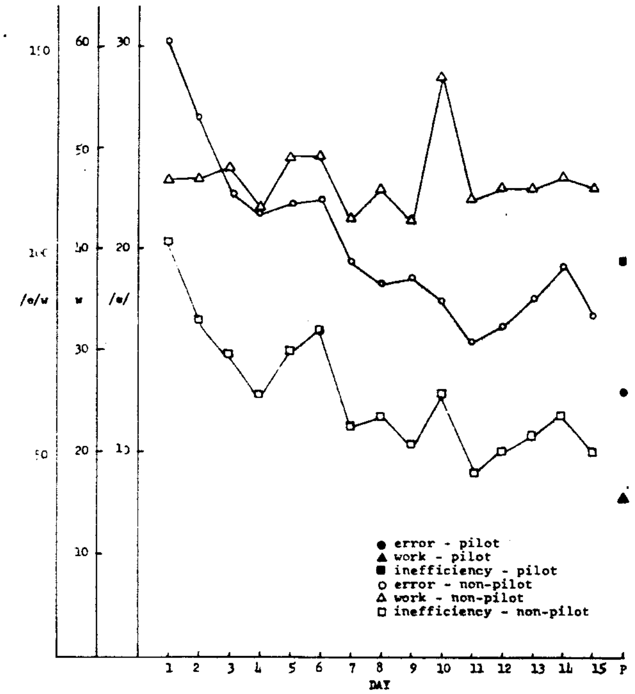

Figure 32 Error (/e/) and inefficiency (/e/w) and work (w) for pitch alone.

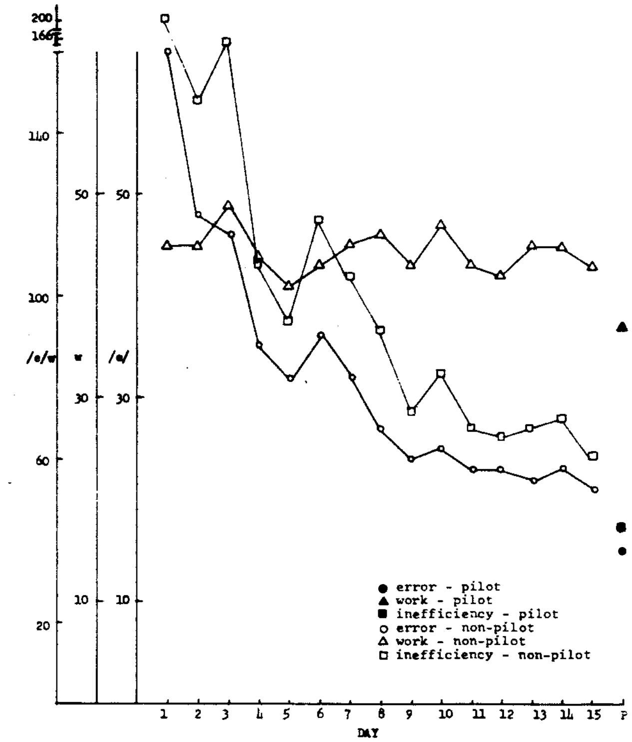

Figure 33 Error (/e/) and inefficiency (/e/w) and work (w) for pitch in the presence of roll.

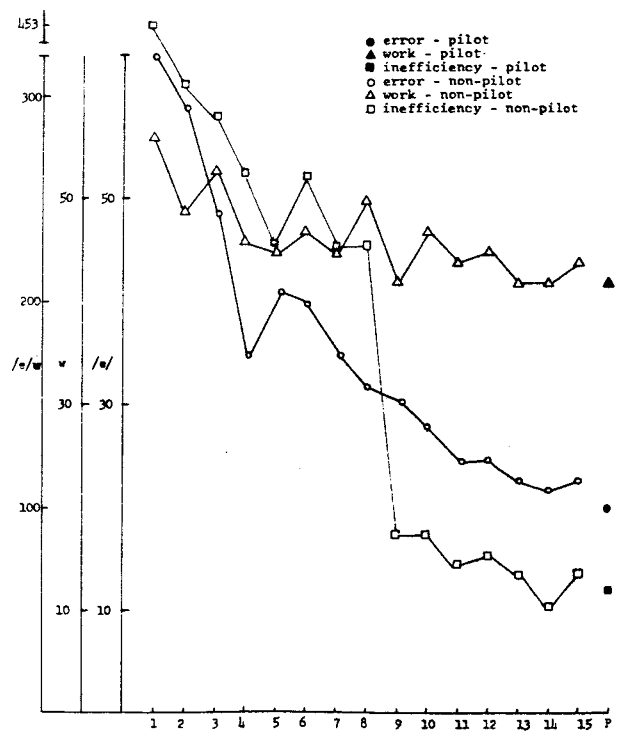

Figure 34 Error (/e/) and inefficiency (/e/w) and work (w) for pitch in the presence of yaw.

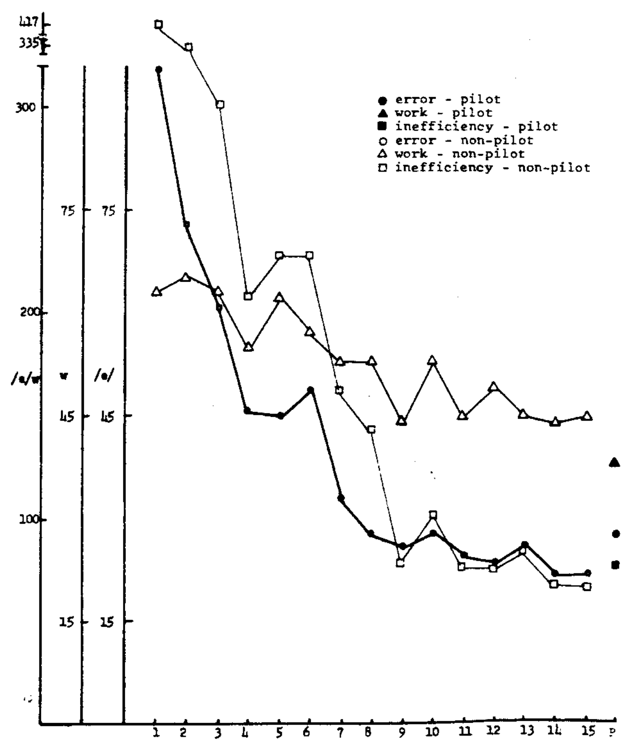

Figure 35 Error (/e/) and inefficiency (/e/w) and work (w) for pitch in the presence of altitude.

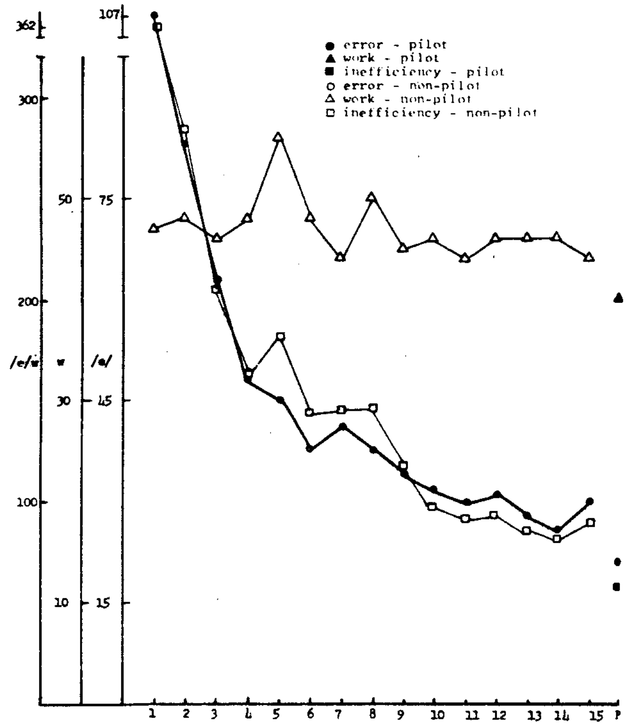

Figure 36 Error (/e/) and inefficiency (/e/w) and work (w) for pitch in the presence of yaw and roll.

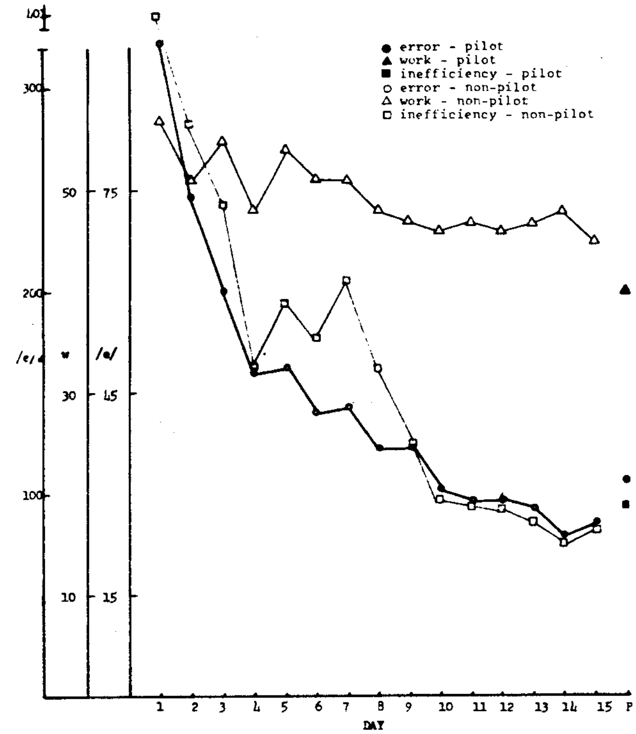

Figure 37 Error (/e/) and inefficiency (/e/w) and work (w) for pitch in the presence of roll and altitude.

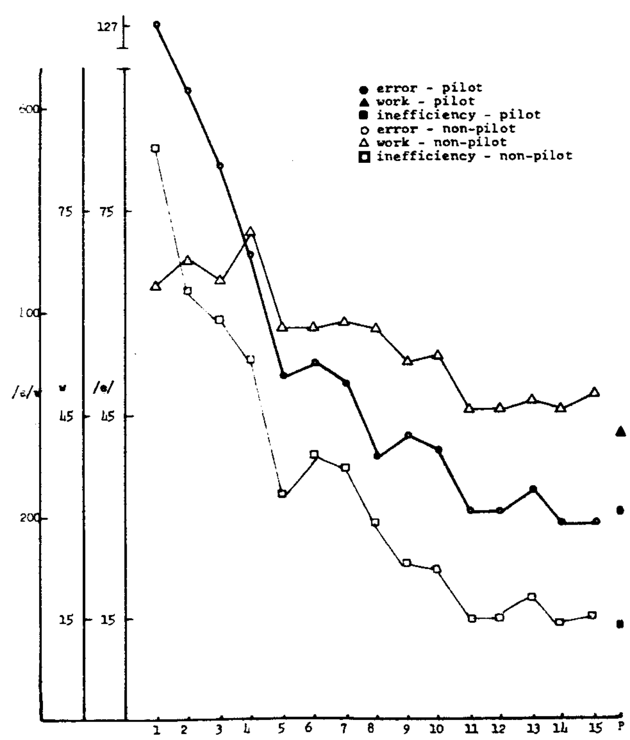

Figure 38 Error (/e/) and inefficiency (/e/w) and work (w) for pitch in the presence of yaw and altitude.

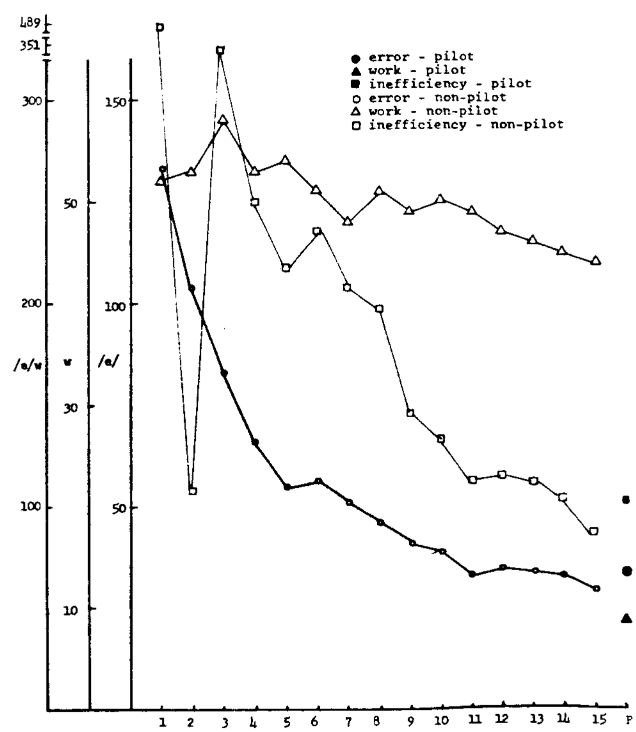

Figure 39 Error (/e/) and inefficiency (/e/w) and work (w) for pitch in the presence of roll, yaw, and altitude.

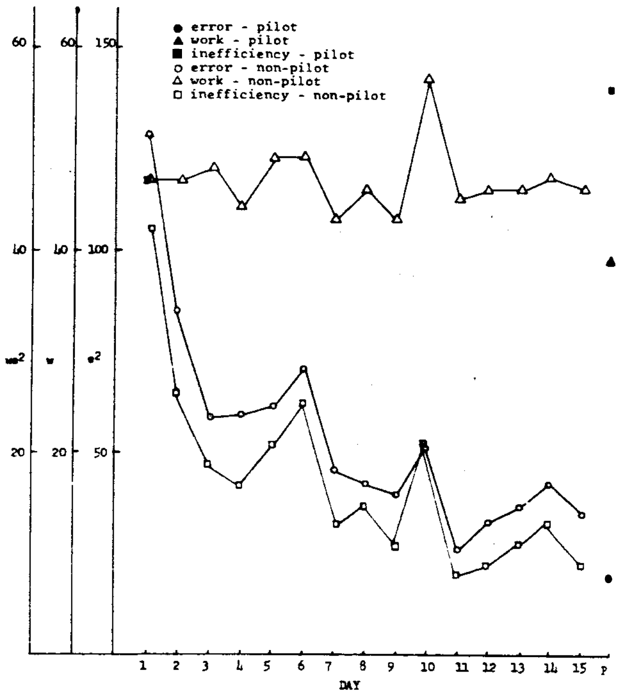

Figure 40 Error (e2) inefficiency (e2w) and work (w) for pitch alone.

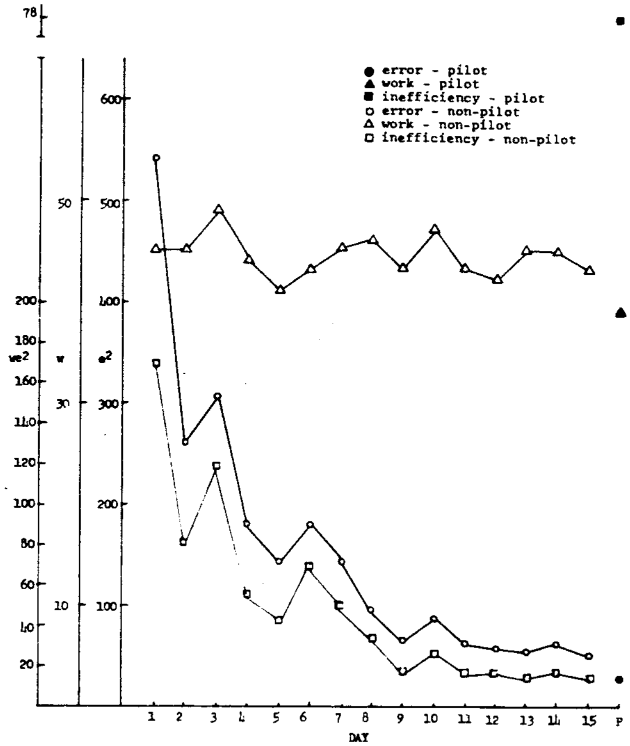

Figure 41 Error (e2) inefficiency (e2w) and work (w) for pitch in the presence of roll.

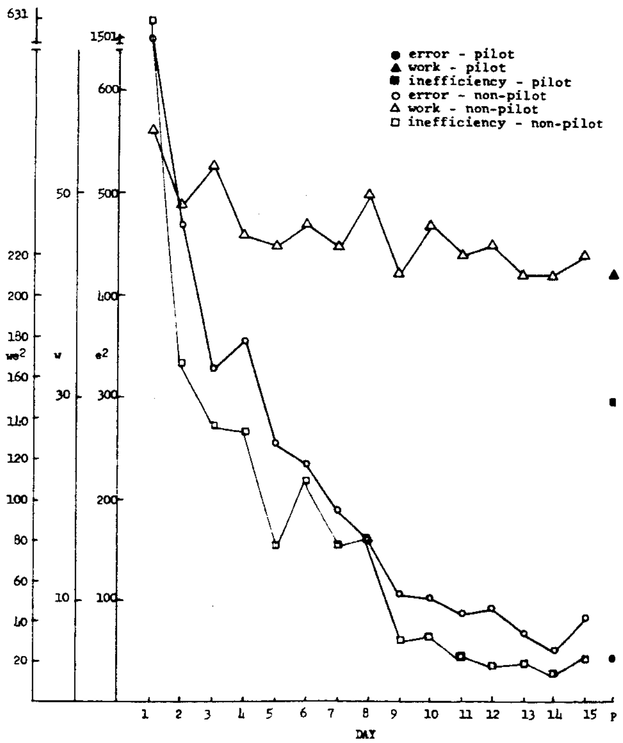

Figure 42 Error (e2) inefficiency (e2w) and work (w) for pitch in the presence of yaw.

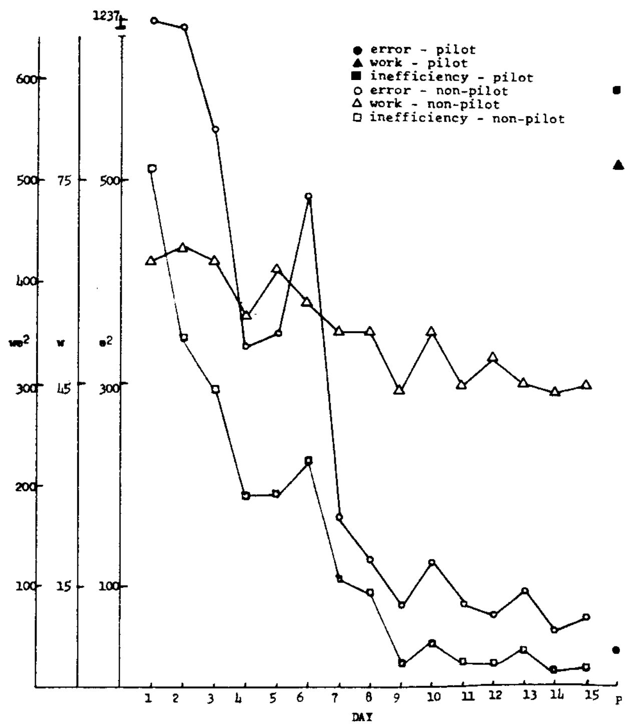

Figure 43 Error (e2) inefficiency (e2w) and work (w) for pitch in the presence of altitude.

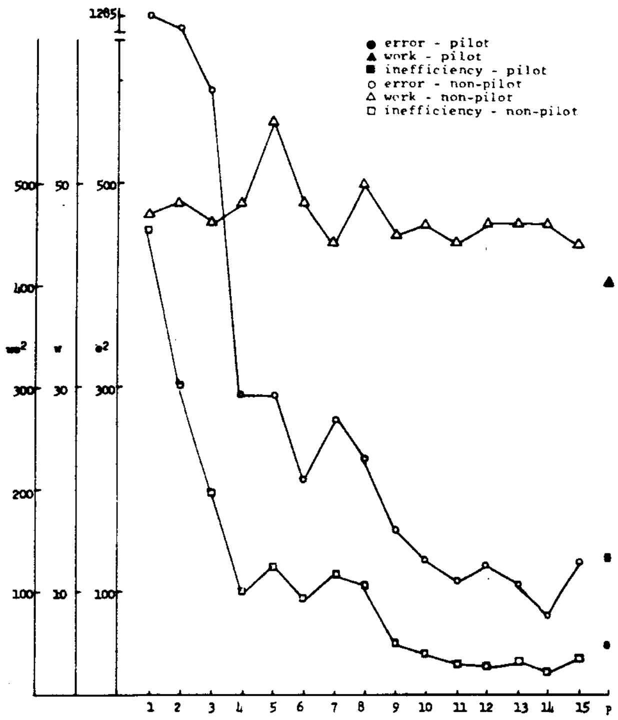

Figure 44 Error (e2) inefficiency (we2) and work (w) for pitch in the presence of roll and yaw.

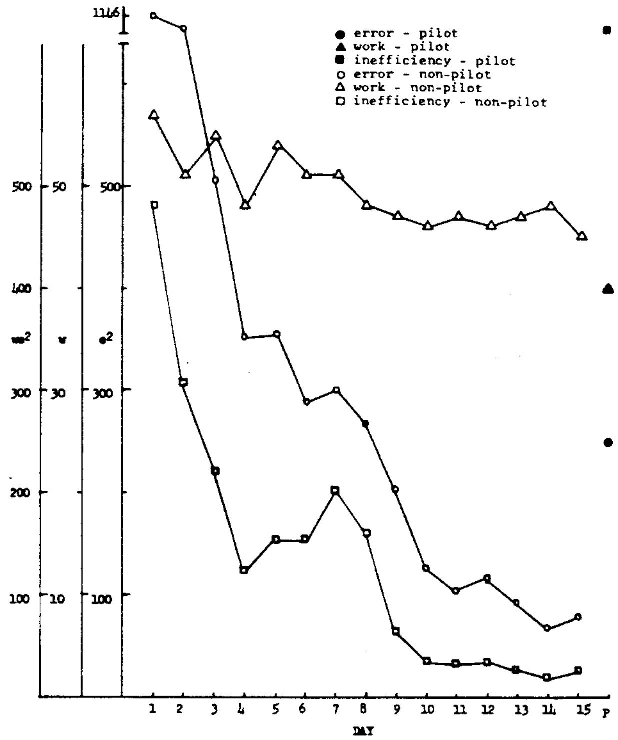

Figure 45 Error(e2) inefficiency (we2) and work (w) for pitch in the presence of roll and altitude.

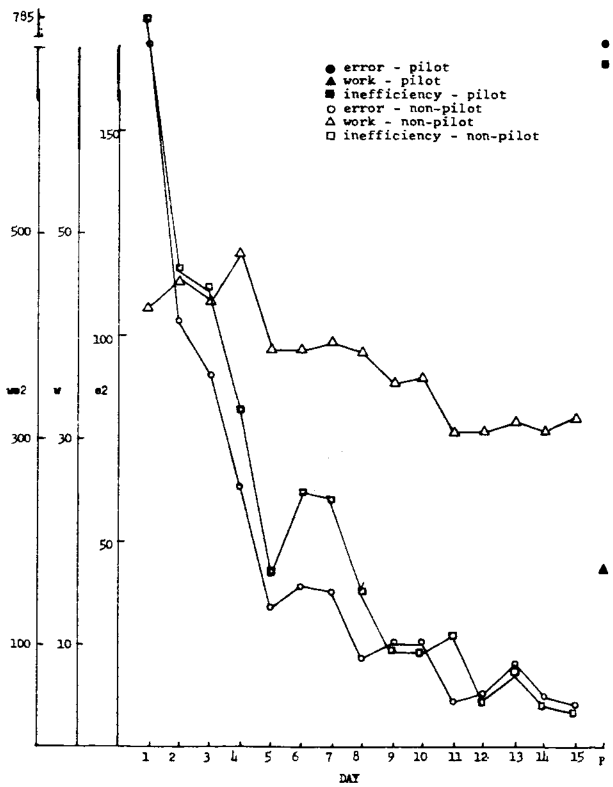

Figure 46 Error (e2) inefficiency (we2) and work (w) for pitch in the presence of yaw and altitude.

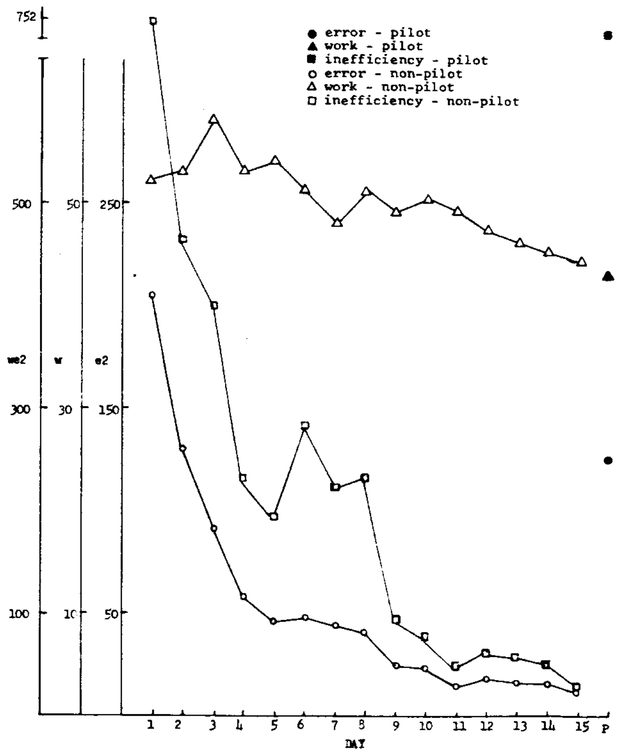

Figure 47 Error (e2) inefficiency (we2) and work (w) for pitch in the presence of roll, yaw and altitude.

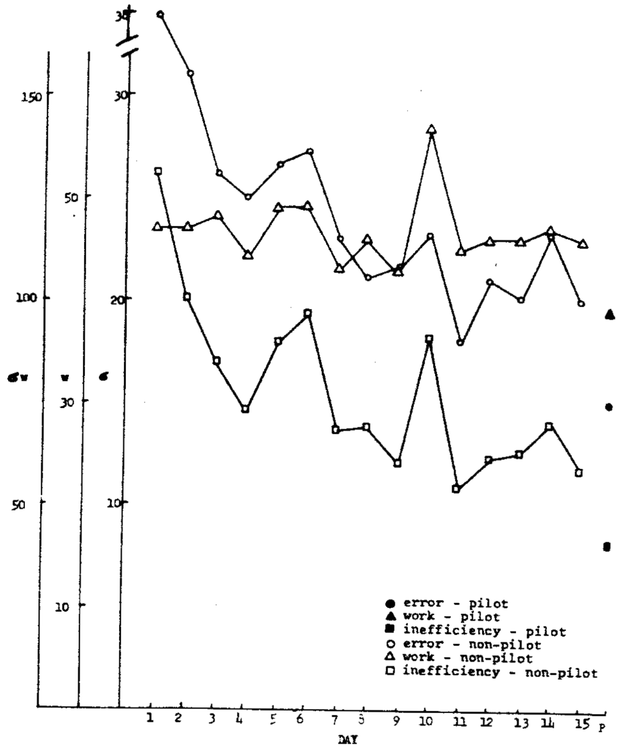

Figure 48 Error (σ) inefficiency (σw) and work (w) for pitch alone.

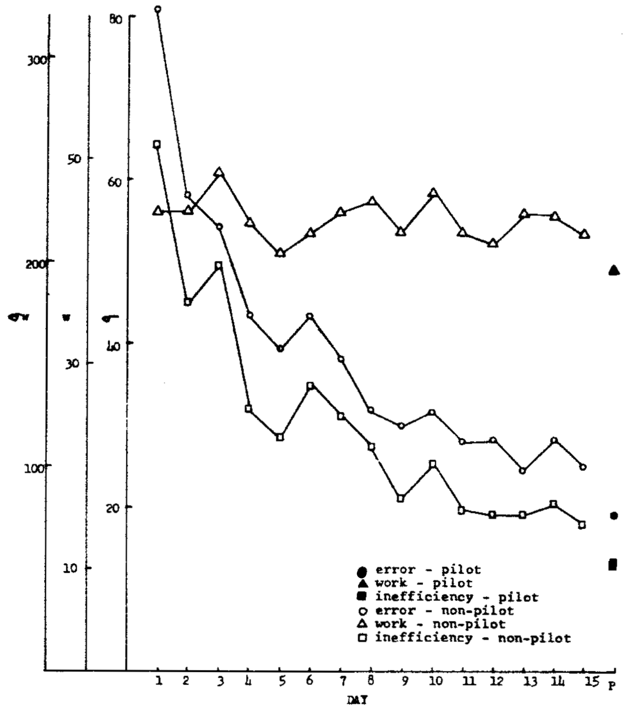

Figure 49 Error (σ) inefficiency (σw) and work (w) for pitch in the presence of roll.

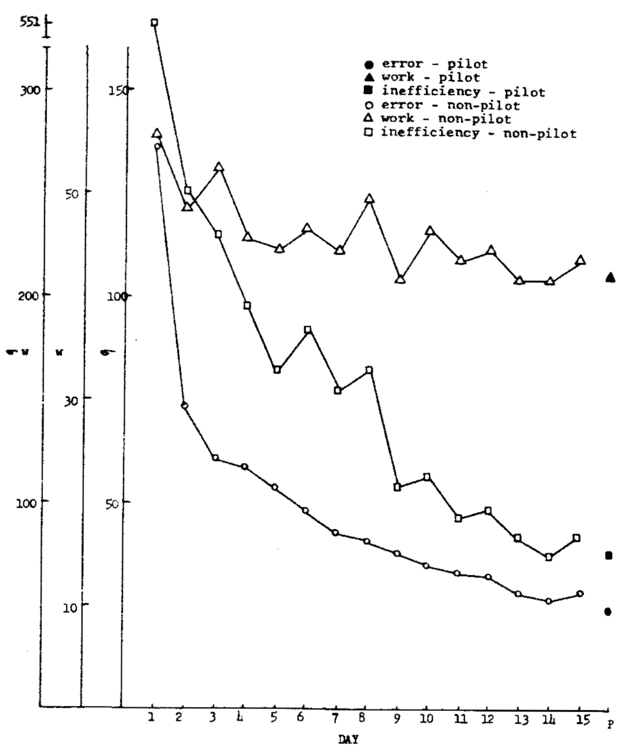

Figure 50 Error (σ) inefficiency (σw) and work (w) for pitch in the presence of yaw.

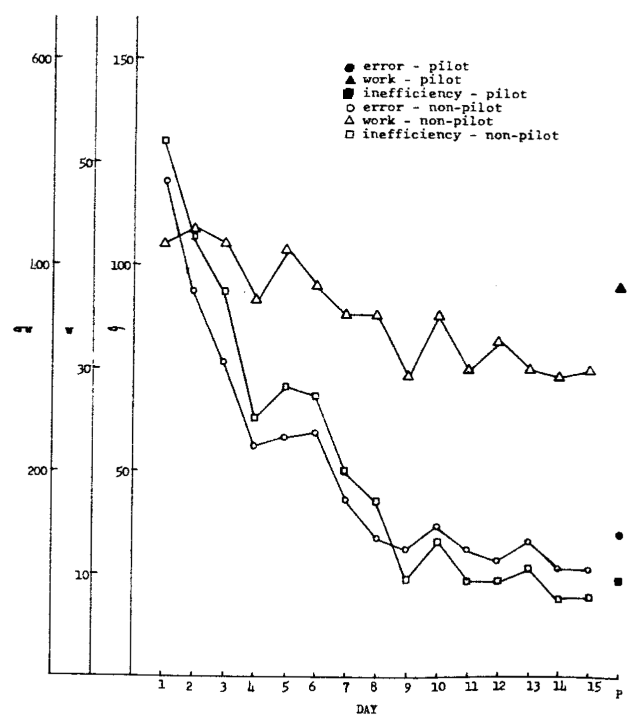

Figure 51 Error (σ) inefficiency (σw) and work (w) for pitch in the presence of altitude.

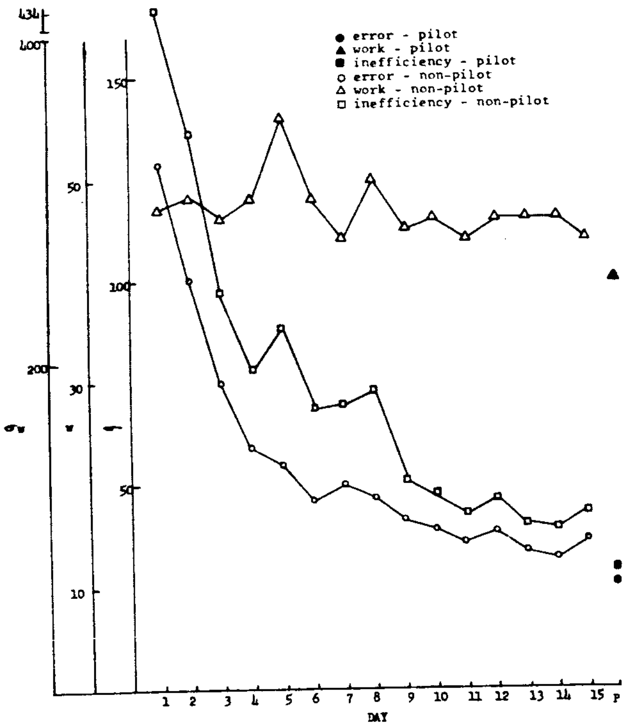

Figure 52 Error (σ) inefficiency (σw) and work (w) for pitch in the presence of roll and yaw.

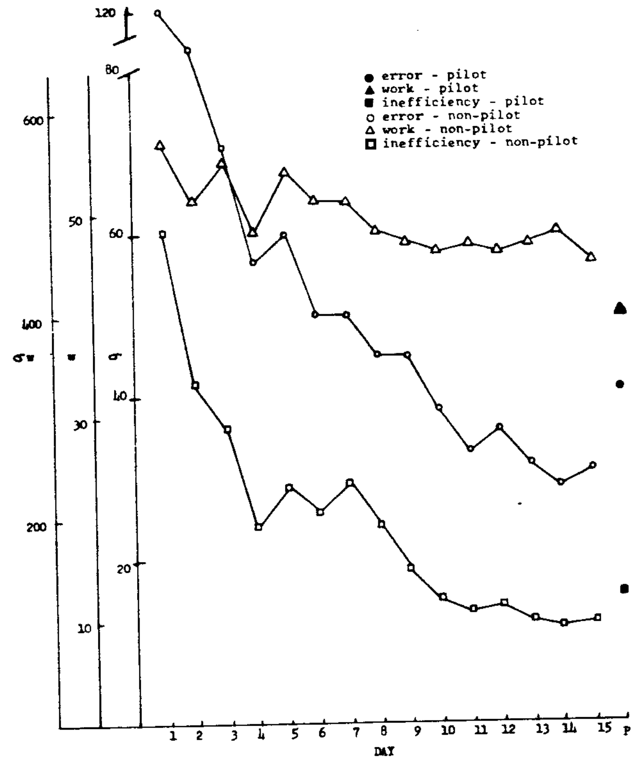

Figure 53 Error (σ) inefficiency (σw) and work (w) for pitch in the presence of roll and altitude.

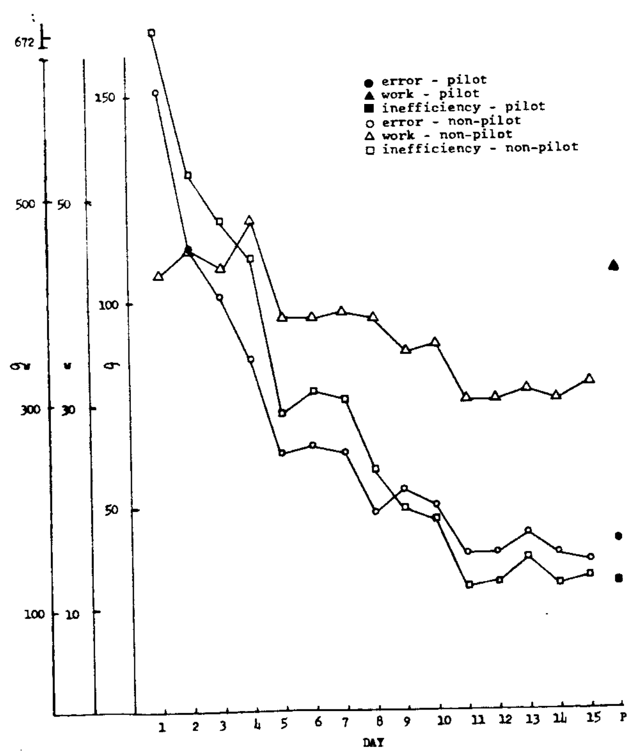

Figure 54 Error (σ) ineffeciency (σw) and work (w) for pitch in the presence of yaw and altitude.

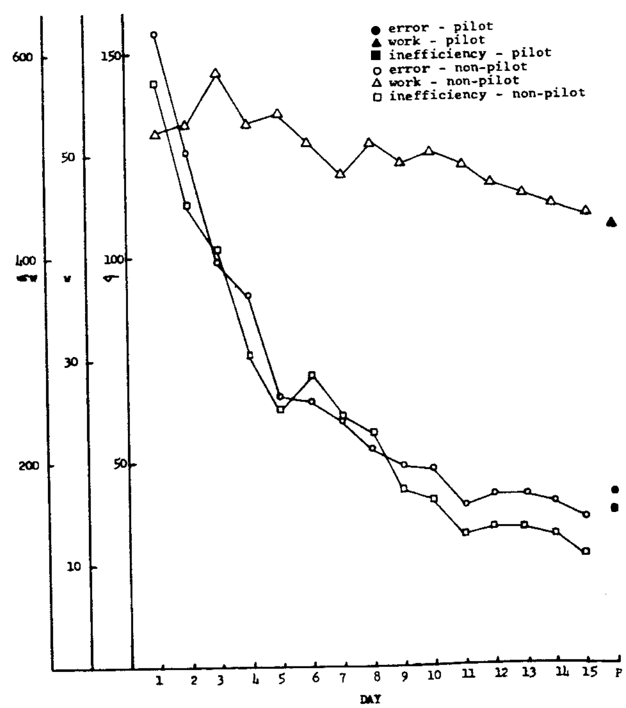

Figure 55 Error (σ) inefficiency (σw) and work (w) for pitch in the presence of roll, yaw and altitude.

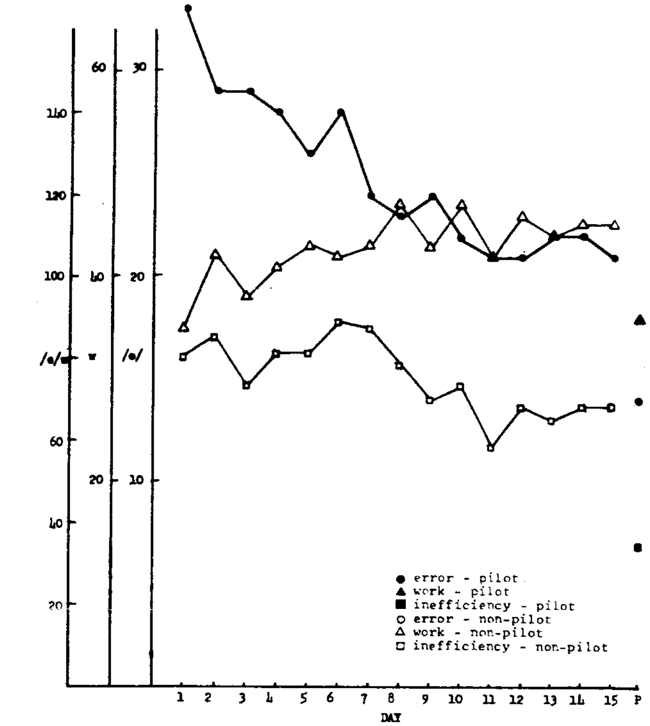

Figure 56 Error (/e/) and inefficiency (/e/w) and work (w) for roll alone.

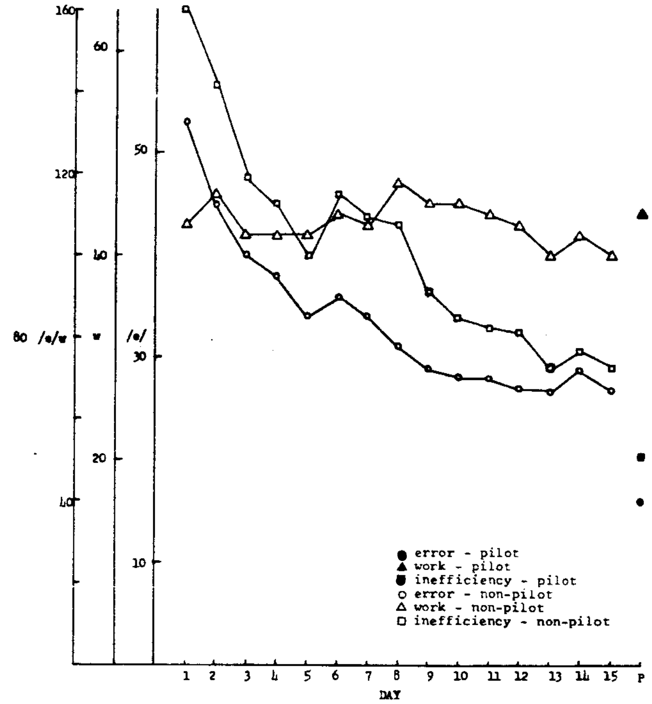

Figure 57 Error (/e/) and inefficiency (/e/w) and work (w) for roll in the presence of pitch.

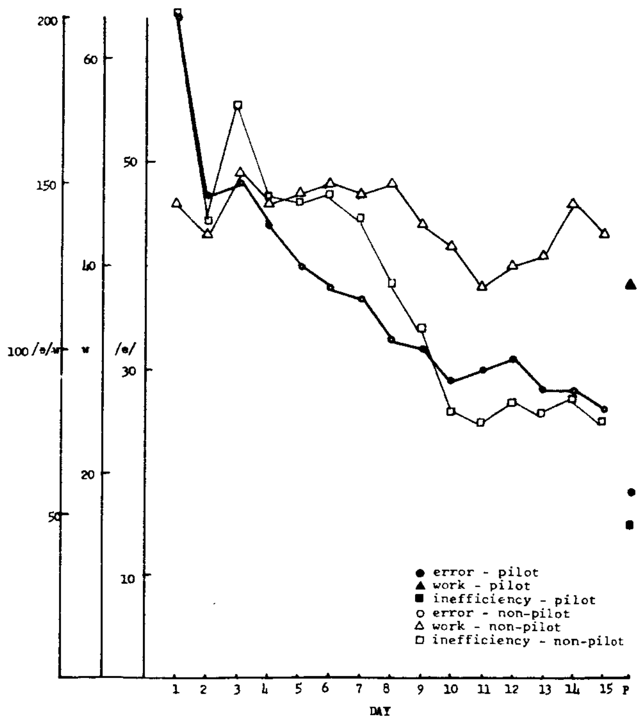

Figure 58 Error (/e/) and inefficiency (/e/w) and work (w) for roll in the presence of yaw.

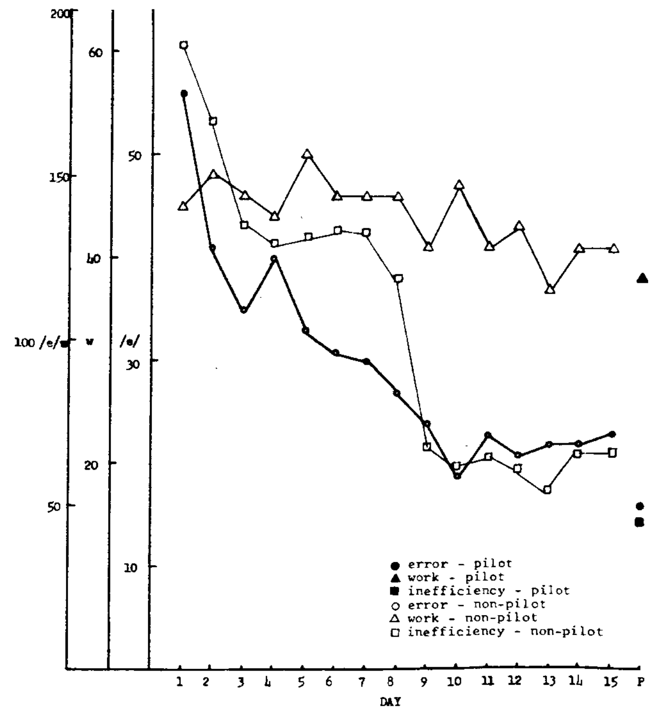

Figure 59 Error (/e/) and inefficiency (/e/w) and work (w) for roll in the presence of altitude.

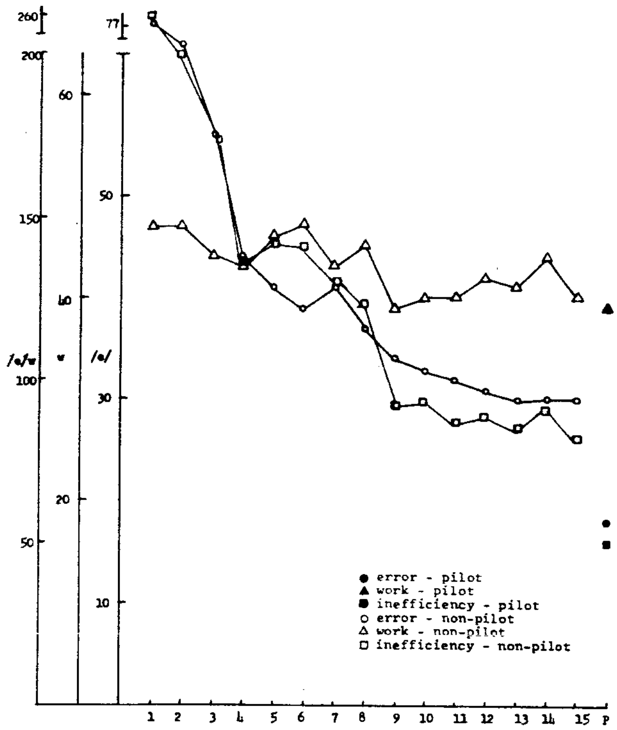

Figure 60 Error (/e/) and inefficiency (/e/w) and work (w) for roll in the presence of pitch and yaw.

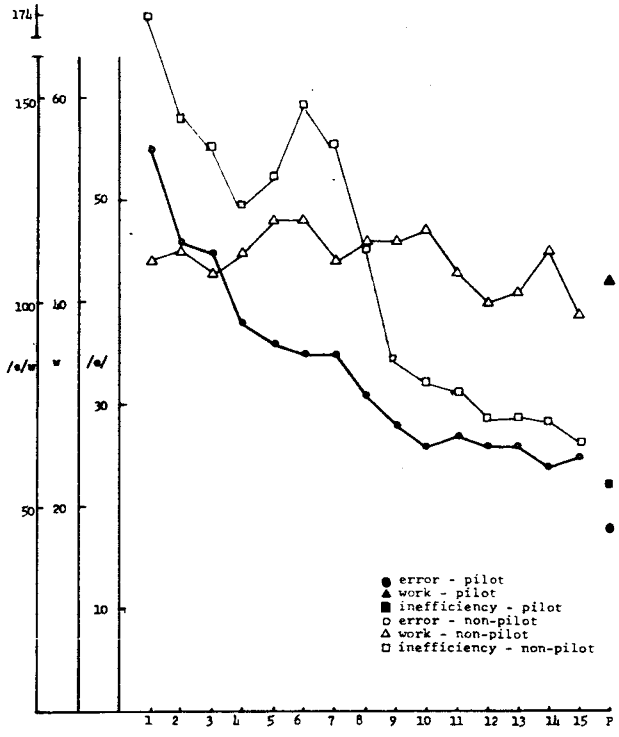

Figure 61 Error (/e/) and inefficiency (/e/w) and work (w) for roll in the presence of pitch and altitude.

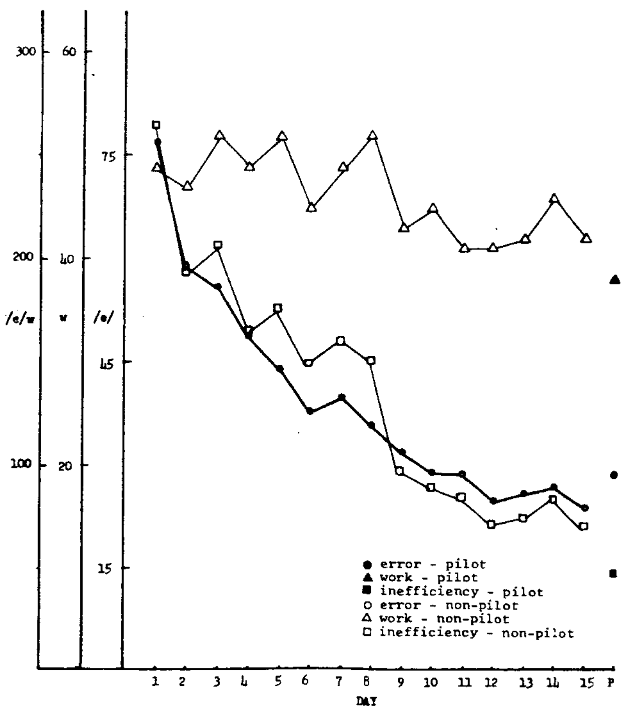

Figure 62 Error (/e/) and inefficiency (/e/w) and work (w) for roll in the presence of yaw and altitude.

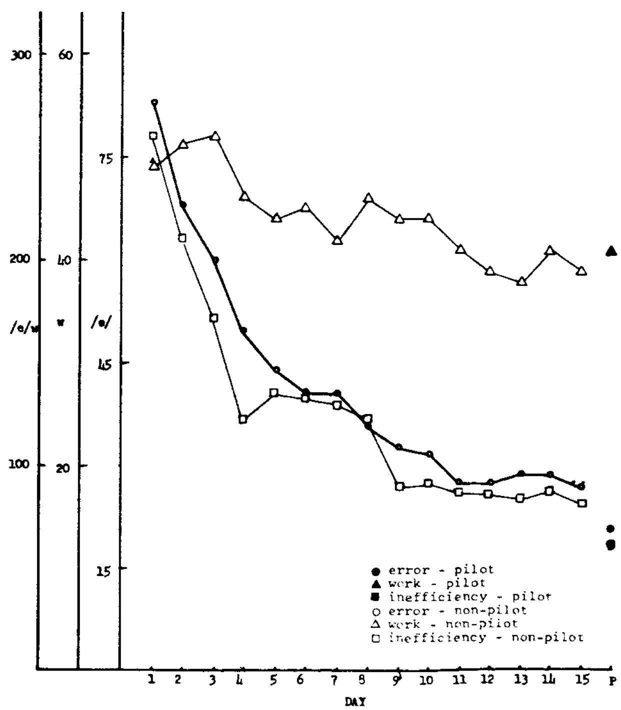

Figure 63 Error (/e/) and inefficiency (/e/w) and work (w) for roll in the presence of pitch, yaw, and altitude.

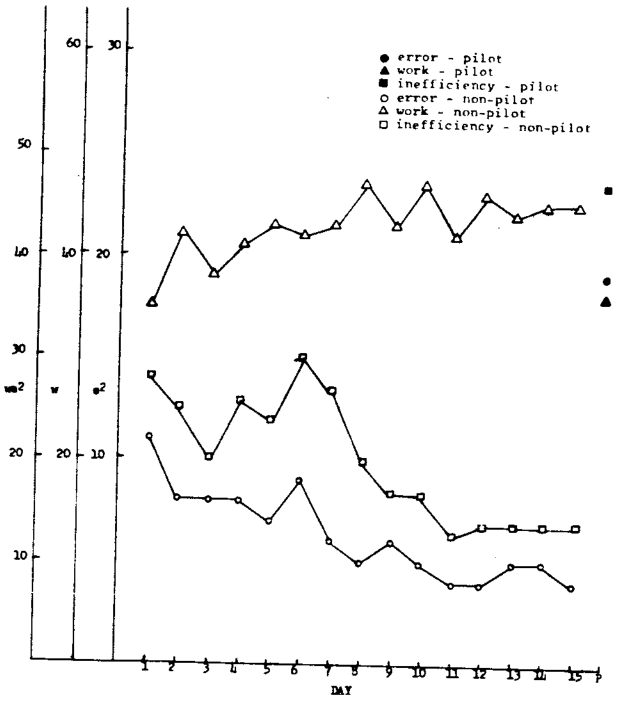

Figure 64 Error (e2) inefficiency (we2) and work (w) for roll alone.

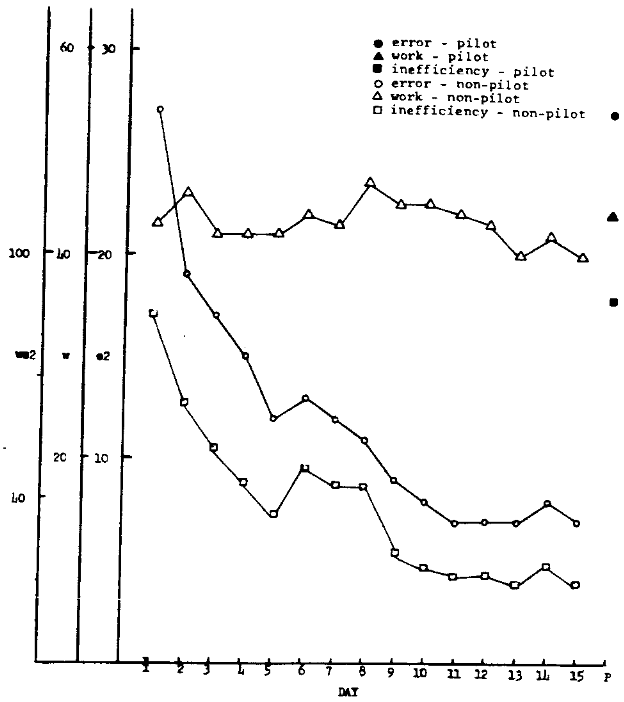

Figure 65 Error (e2) inefficiency (we2) and work (w) for roll in the presence of pitch.

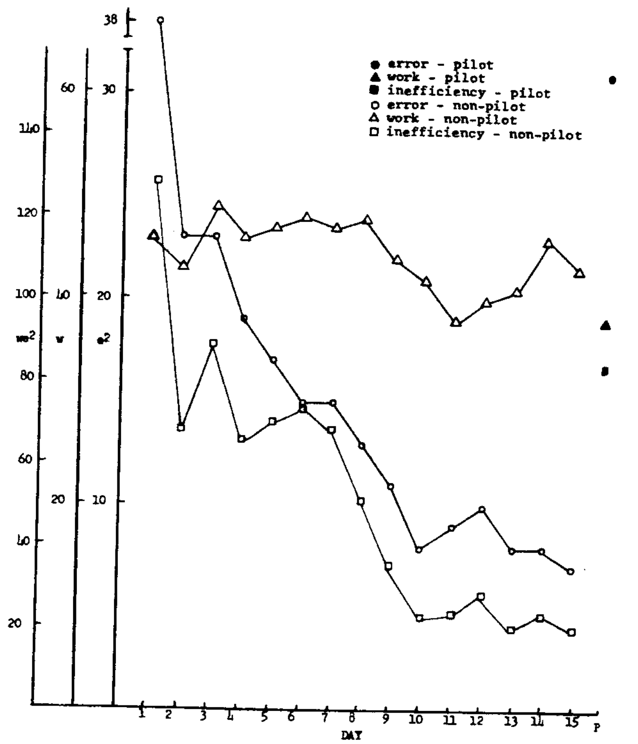

Figure 66 Error (e2) inefficiency (we2) and work (w) for roll in the presence of yaw.

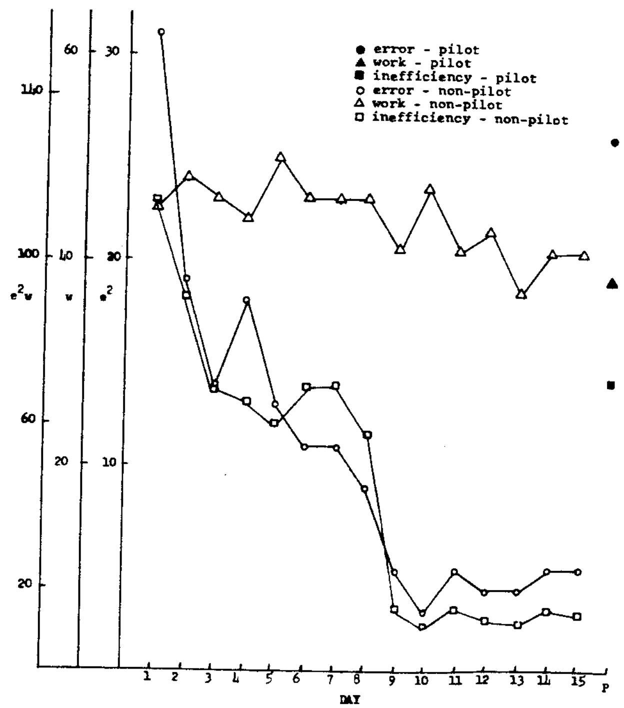

Figure 67 Error (e2) inefficiency (e2w) and work (w) for roll in the presence of altitude.

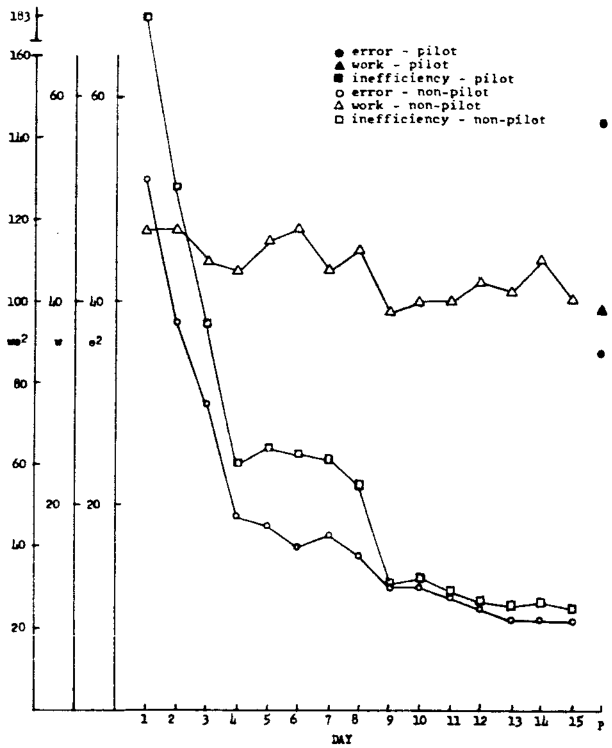

Figure 68 Error (e2) inefficiency (we2) and work (w) for roll in the presence of pitch and yaw.

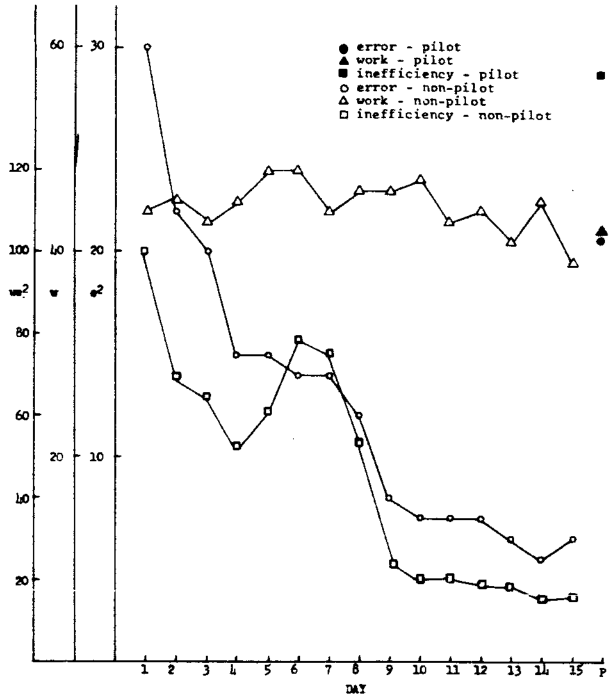

Figure 69 Error (e2) inefficiency (we2) and work (w) for roll in the presence of pitch and altitude.

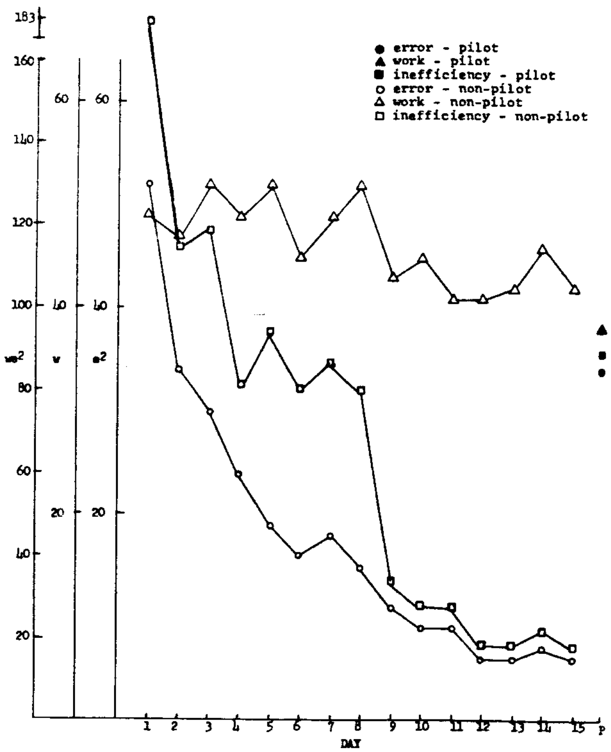

Figure 70 Error (e2) inefficiency (we2) and work (w) for roll in the presence of yaw and altitude.

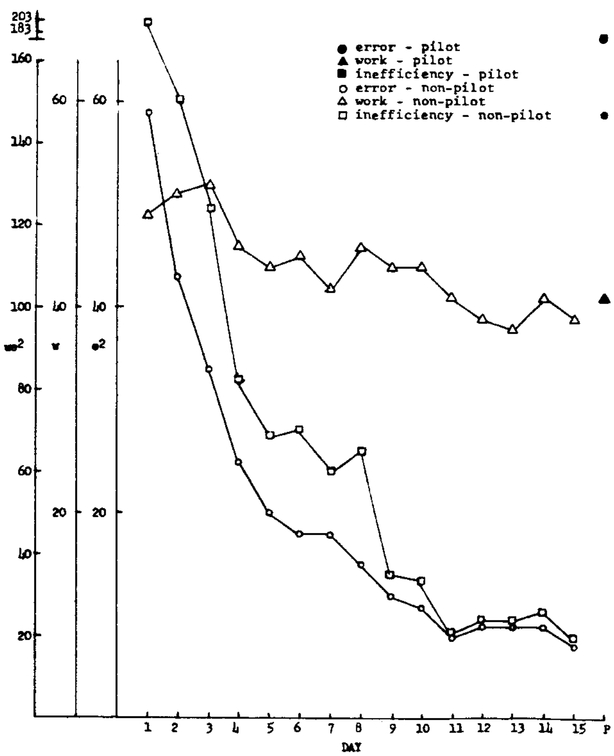

Figure 71 Error (e2) inefficiency (we2) and work (w) for roll in the presence of pitch, yaw and altitude.

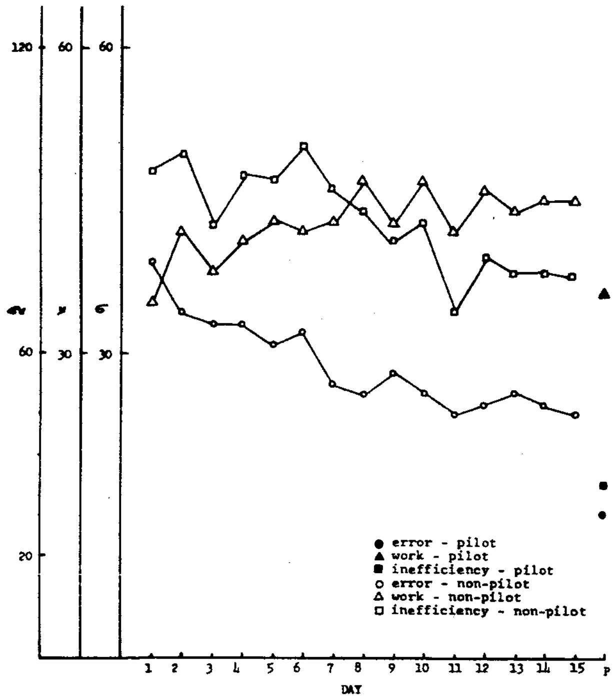

Figure 72 Error (σ) inefficiency (σw) and work (w) for roll alone.

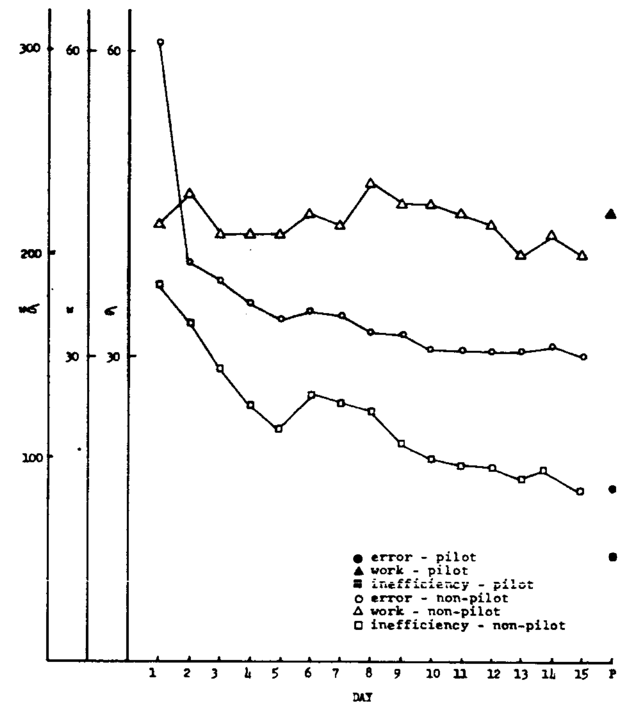

Figure 73 Error (σ) inefficiency (wσ) and work (w) for roll in the presence of pitch.

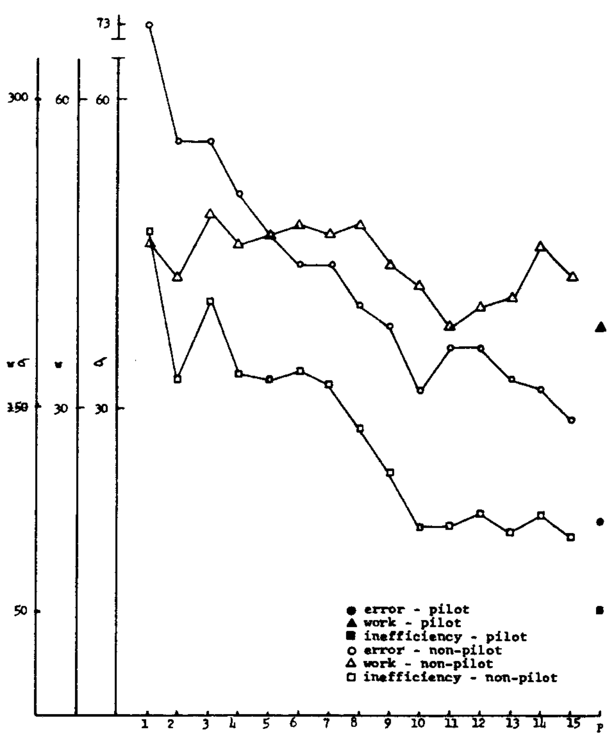

Figure 74 Error (σ) inefficiency (wσ) and work (w) for roll in the presence of yaw.

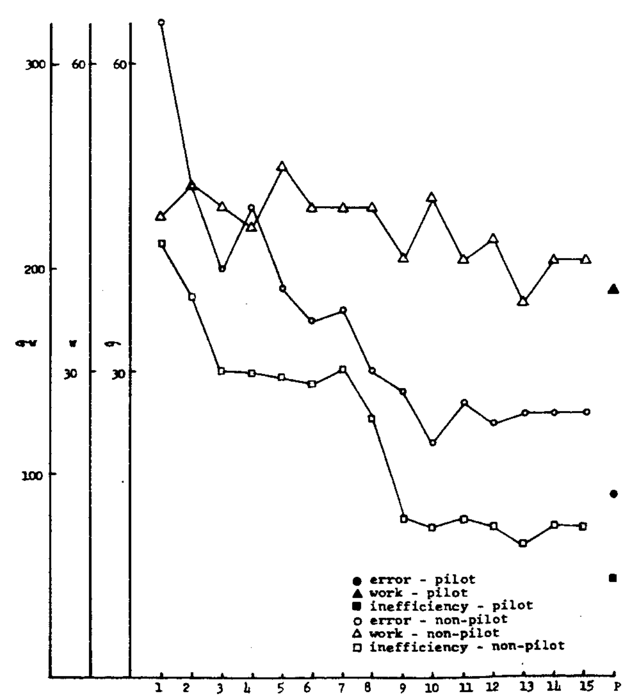

Figure 75 Error (σ) inefficiency (σw) and work (w) for roll in the presence of altitude.

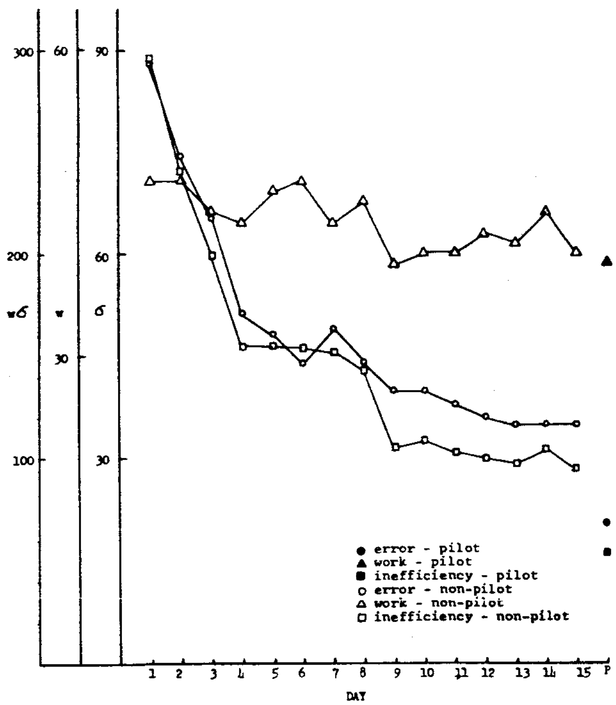

Figure 76 Error (σ) inefficiency (wσ) and work (w) for roll in the presence of roll and yaw.

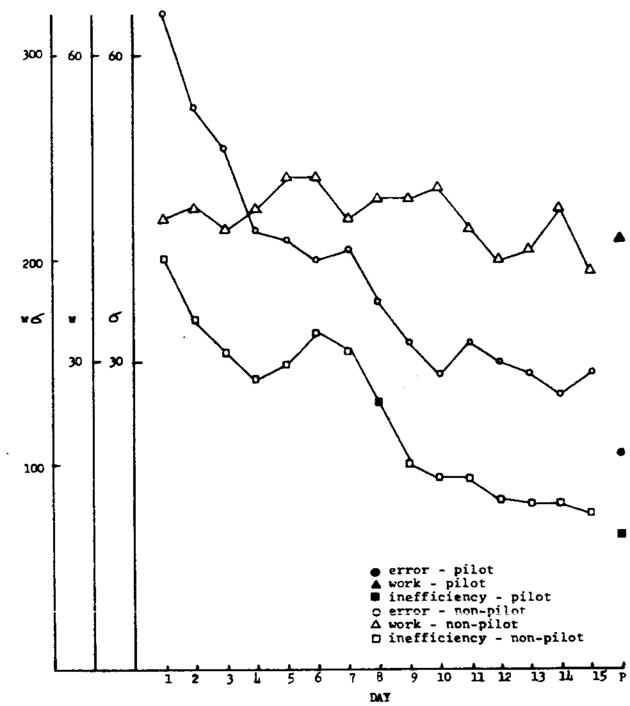

Figure 77 Error (σ) inefficiency (wσ) and work (w) for roll in the presence of pitch and altitude.

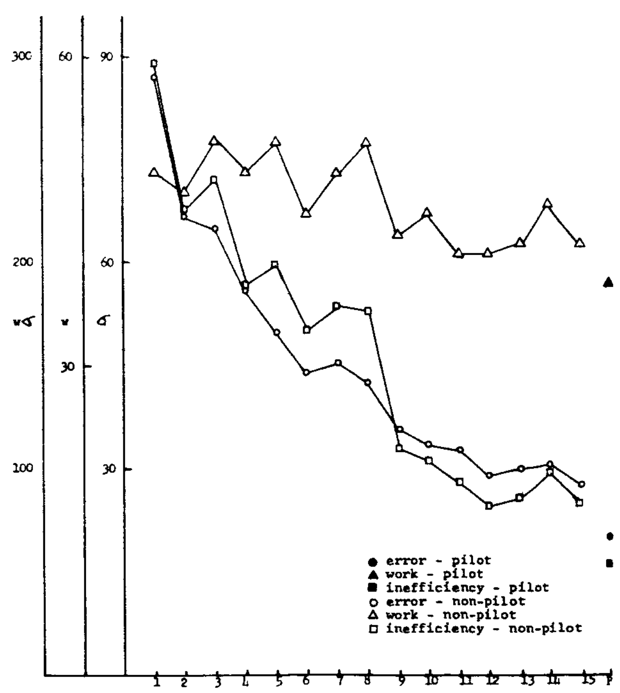

Figure 78 Error (σ) inefficiency (wσ) and work (w) for roll in the presence of yaw and altitude.

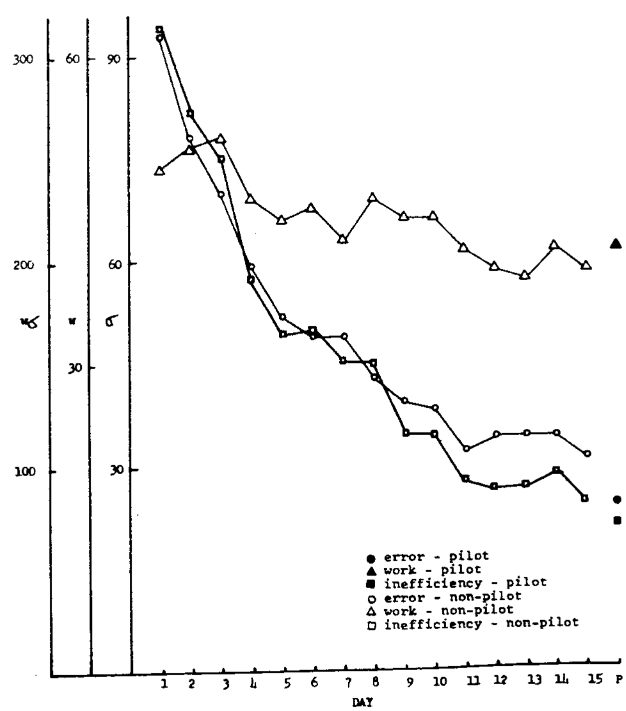

Figure 79 Error (σ) inefficiency (wσ) and work (w) for roll in the presence of pitch, yaw and altitude.

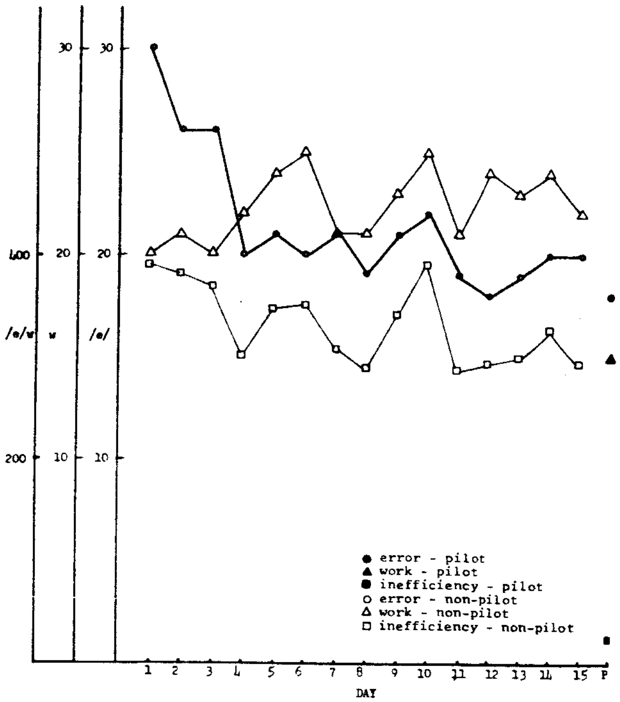

Figure 50 Error (/e/) and inefficiency (/e/w) and work (w) for yaw alone.

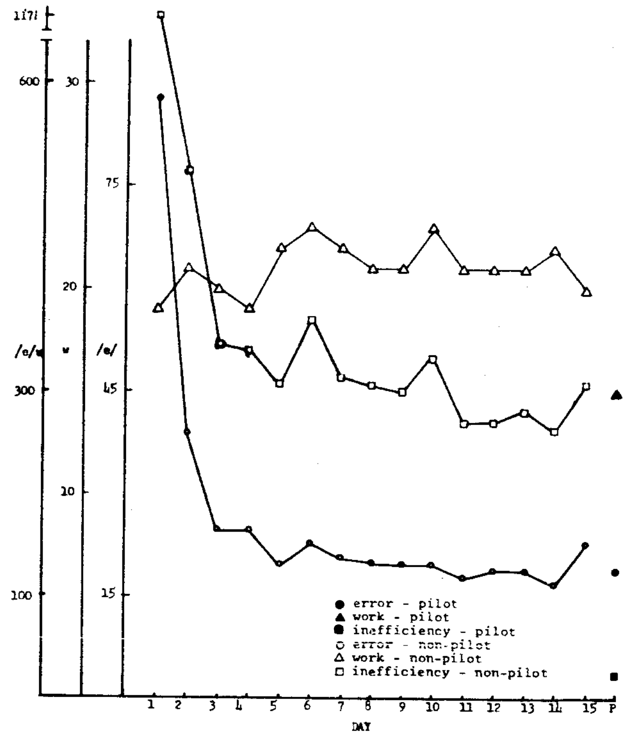

Figure 51 Error (/e/) and inefficiency (/e/w) and work (w) for yaw in the presence of pitch.

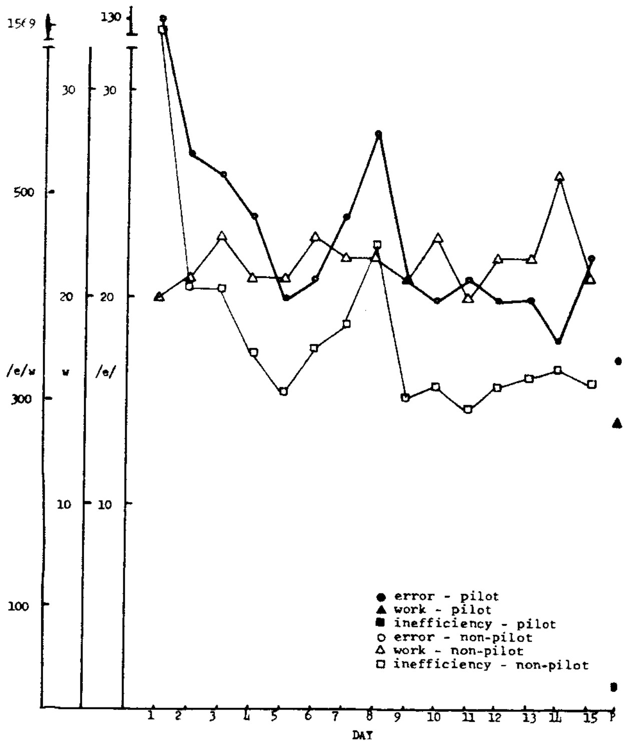

Figure 52 Error (/e/) and inefficiency (/e/w) and work (w) for yaw in the presence of pitch.

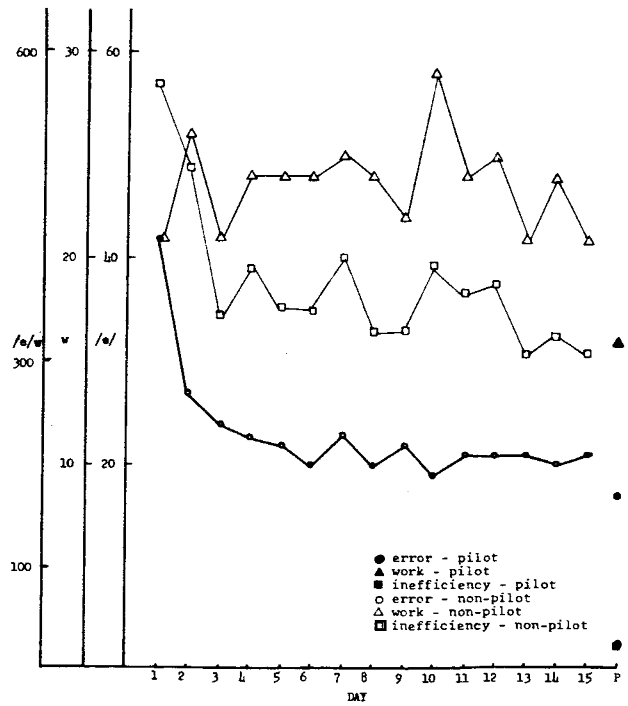

Figure 53 Error (/e/) and inefficiency (/e/w) and work (w) for yaw in the presence of altitude.

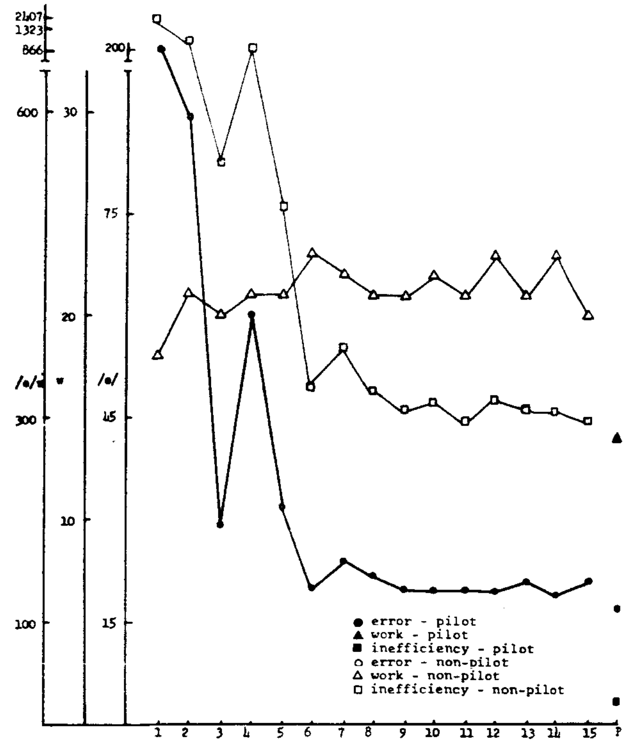

Figure 84 Error (/e/) and inefficiency (/e/w) and work (w) for yaw in the presence of pitch and roll.

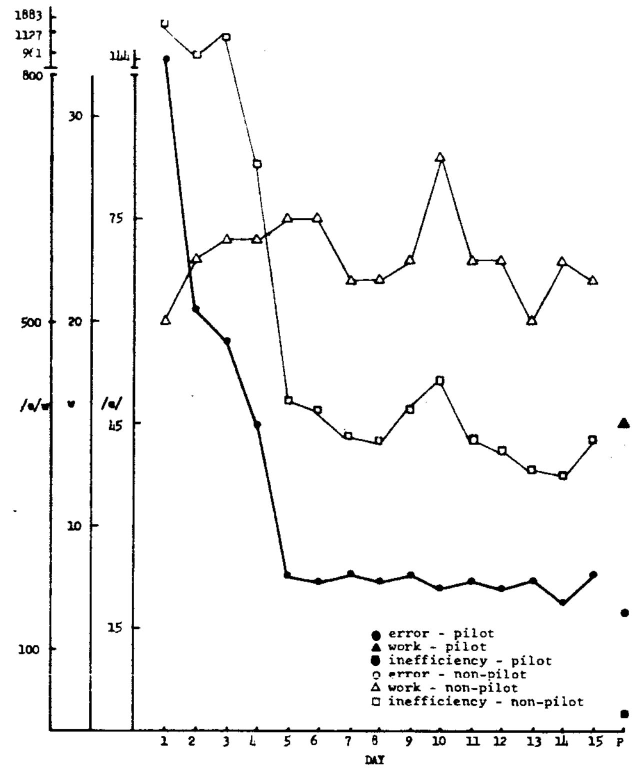

Figure 85 Error (/e/) and inefficiency (/e/w) and work (w) for yaw in the presence of pitch and altitude.

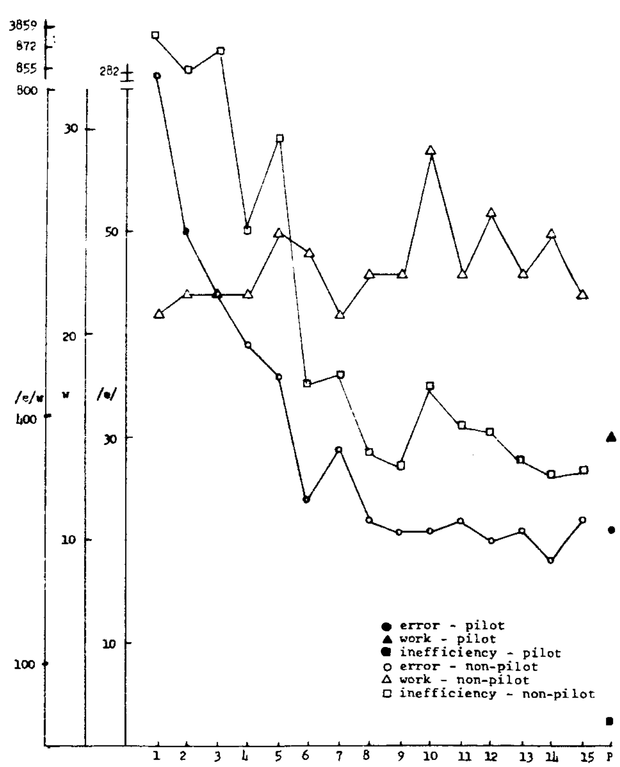

Figure 86 Error (/e/) and inefficiency (/e/w) and work (w) for yaw in the presence of roll and altitude.

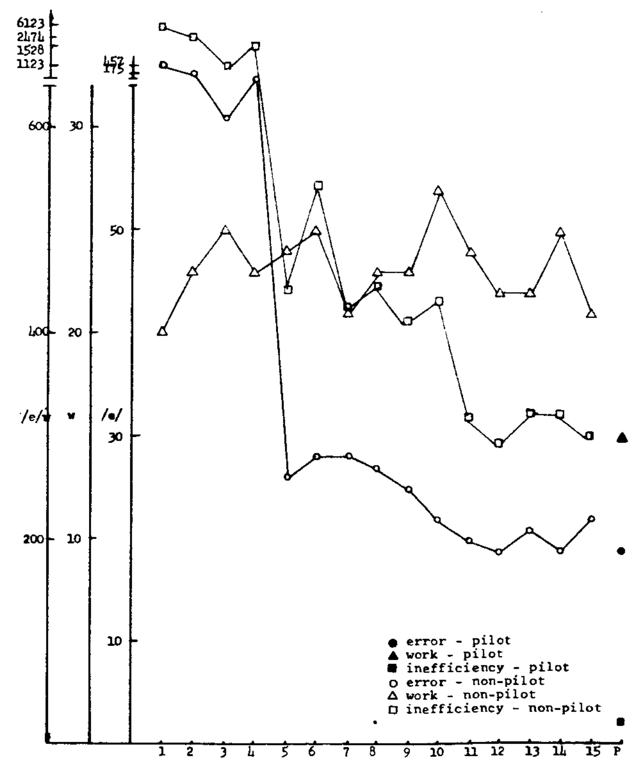

Figure 87 Error (/e/) and inefficiency (/e/w) and work (w) for yaw in the presence of pitch, roll, and altitude.

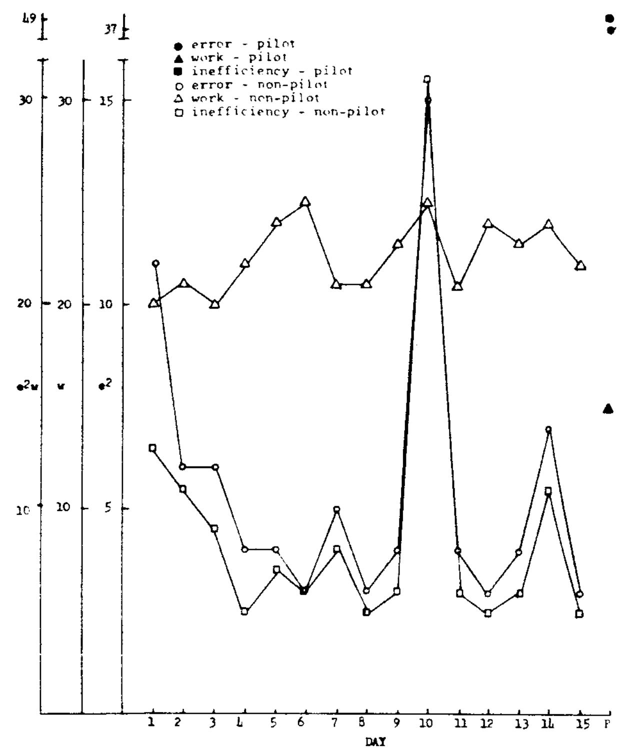

Figure 88 Error (e2) inefficinecy (e2w) and work (w) for yaw alone.

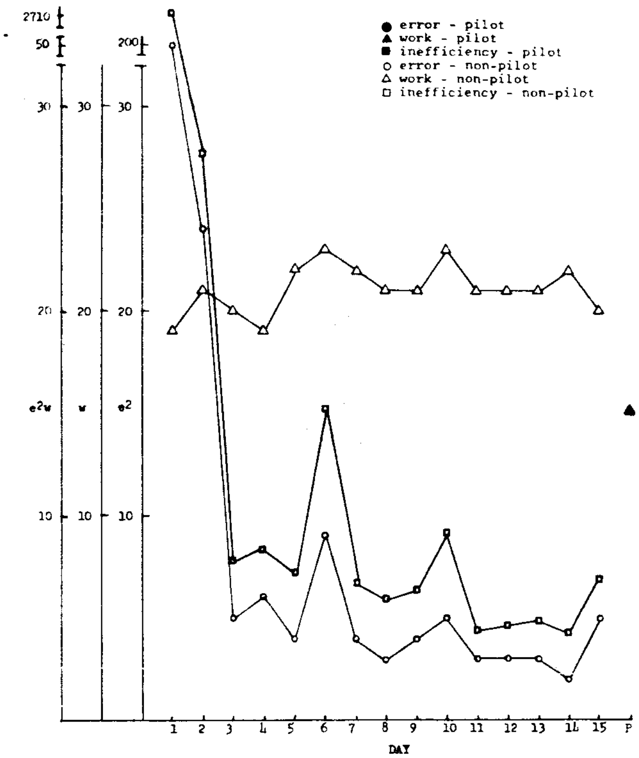

Figure 89 Error (e2) inefficincy (e2w) and work (w) for yaw in the presence of pitch.

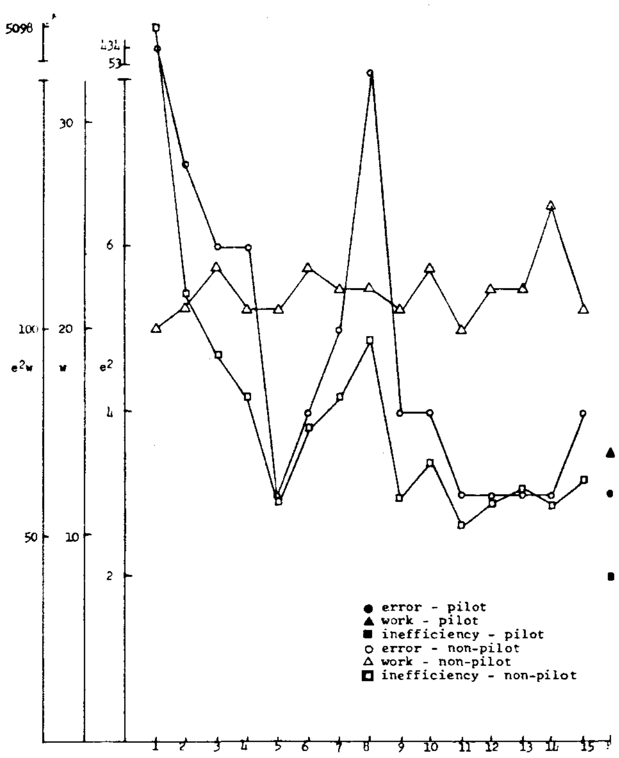

Figure 90 Error (e2) inefficiency (e2w) and work (w) for yaw in the presence of roll.

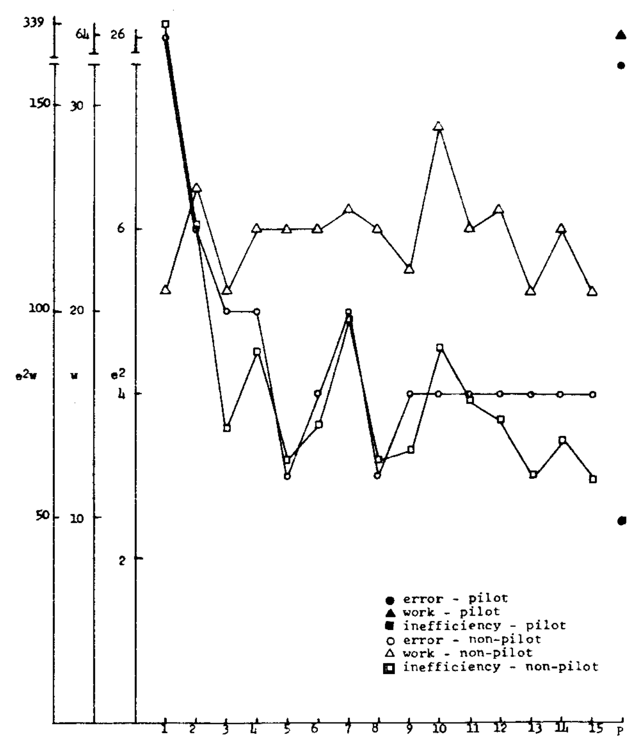

Figure 91 Error (e2) inefficiency (e2w) and work (w) for yaw in the presence of altitude.

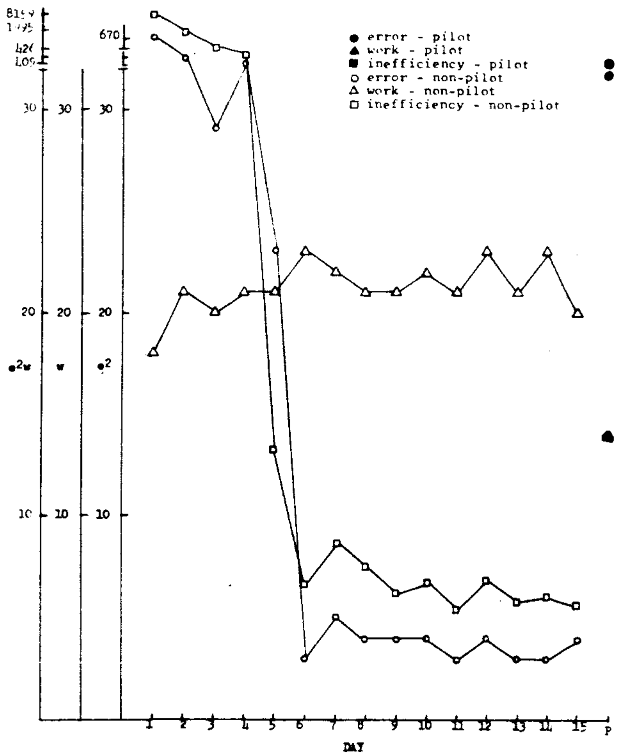

Figure 92 Error (e2) inefficiency (e2w) and work (w) for yaw in the presence of pitch and roll.

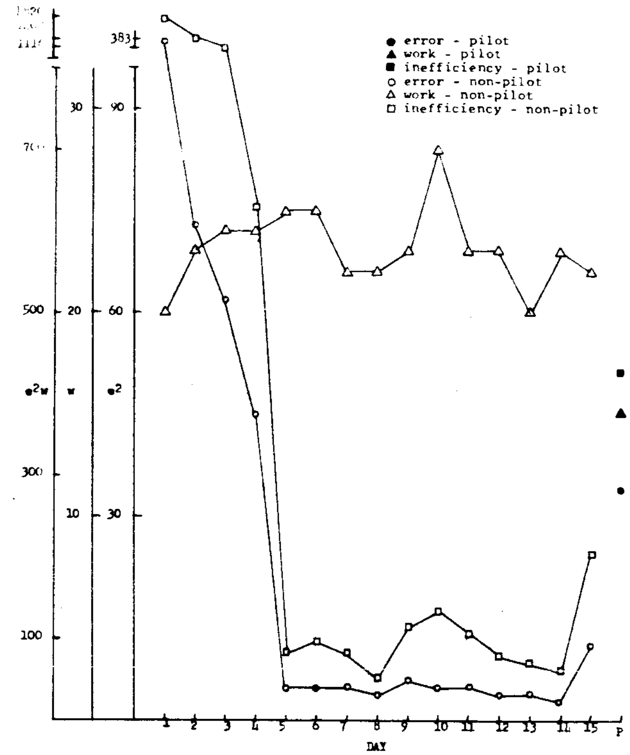

Figure 93 Error (e2) inefficiency (e2w) and work (w) for yaw in the presence of pitch and altitude.

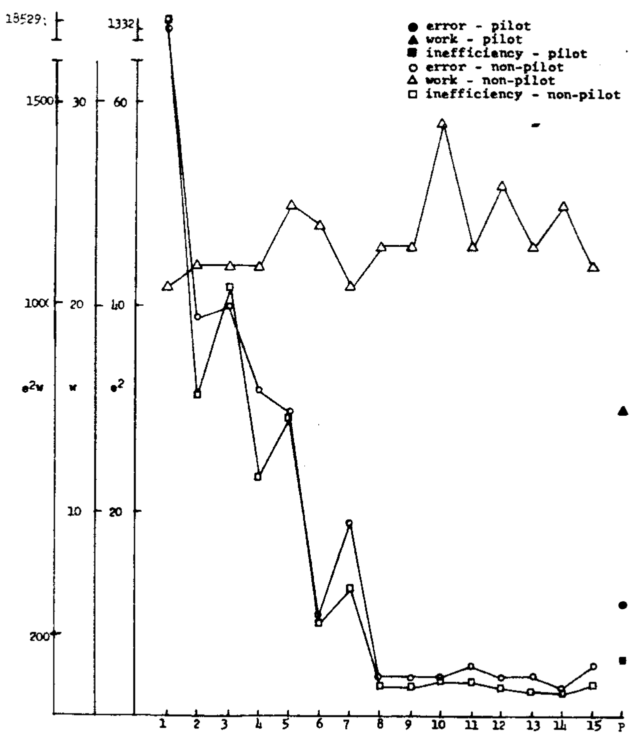

Figure 94 Error (e2) inefficiency (e2w) and work (w) for yaw in the presence of roll and altitude.

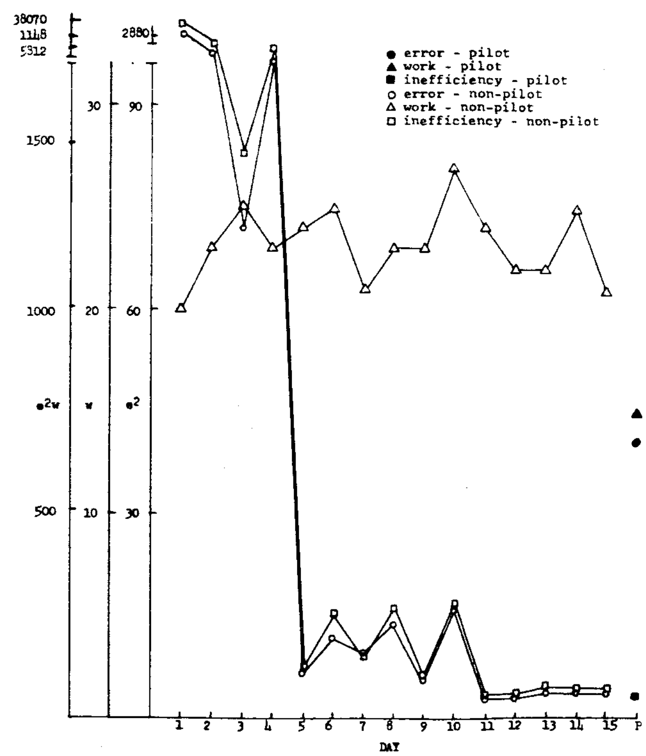

Figure 95 Error (e2) inefficiency (e2w) and work (w) for yaw in the presence of pitch, roll and altitude.

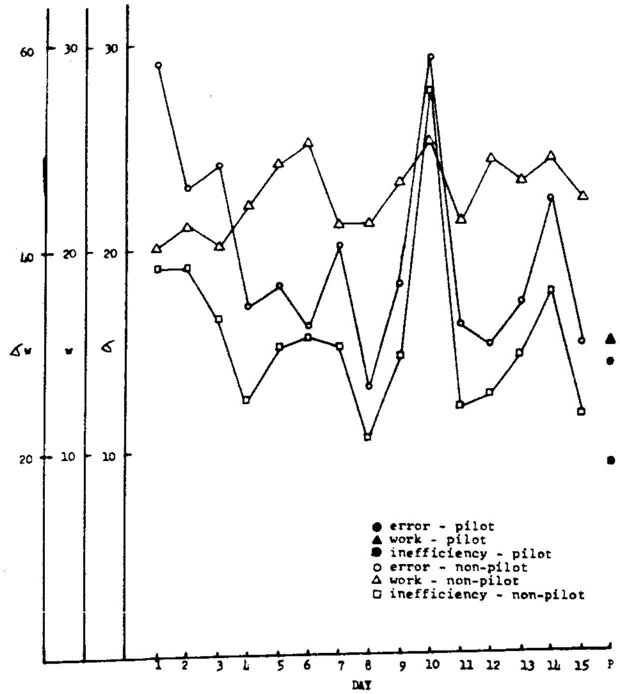

Figure 96 Error (σ) inefficiency (σ_w) and work (w) for yaw alone.

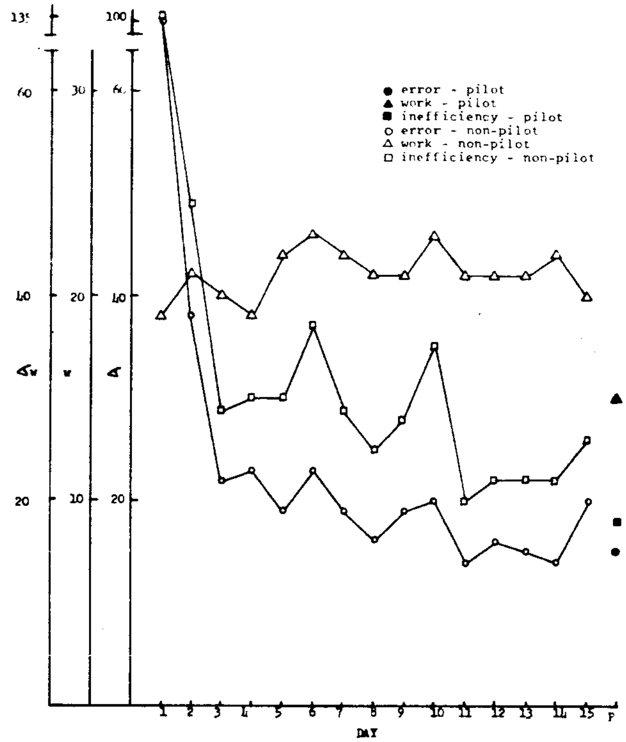

Figure 97 Error (σ) inefficiency (σ_w) and work (w) for yaw in the presence of pitch.

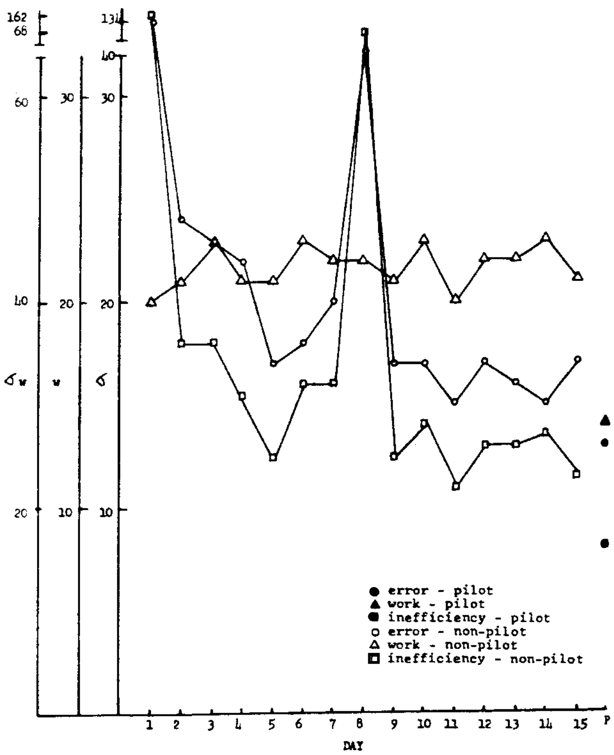

Figure 98 Error (σ) inefficiency (σ_w) and work (w) for yaw in the presence of roll.

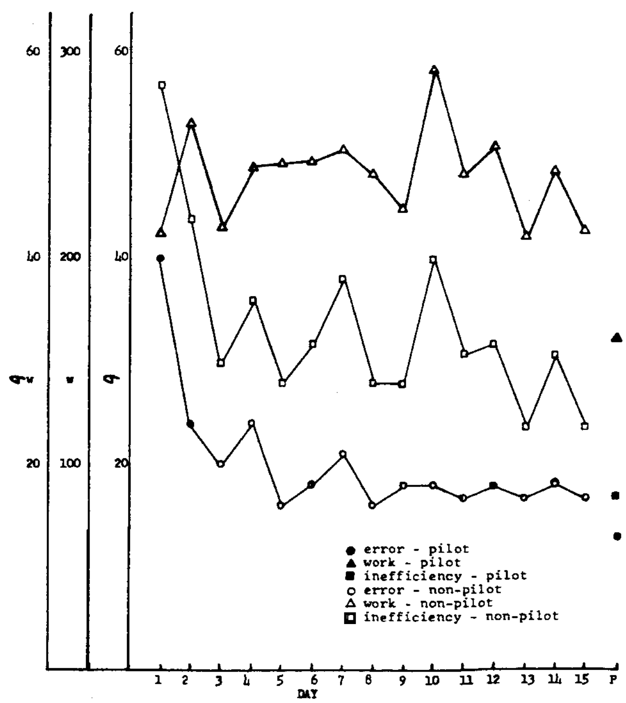

Figure 99 Error (σ) and inefficiency (σ_w) and work (w) for yaw in the presence of altitude.

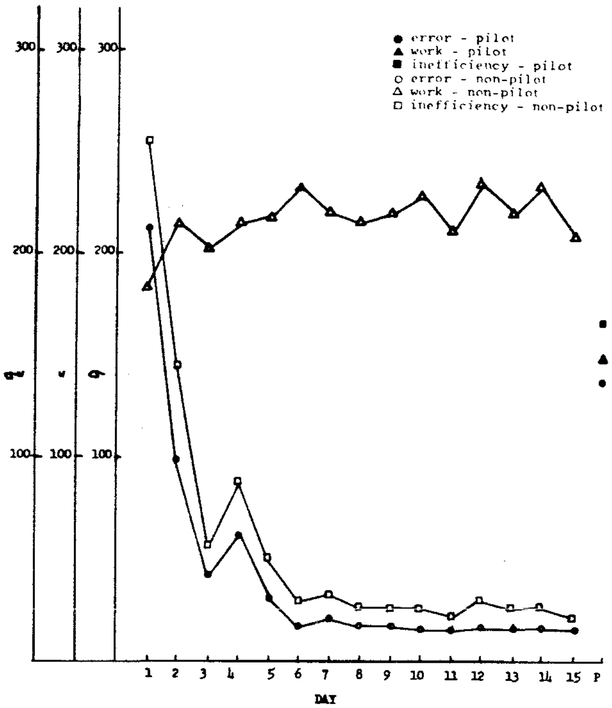

Figure 100 Error (σ) and inefficiency (σw) and work (w) for yaw in the presence of pitch and roll.

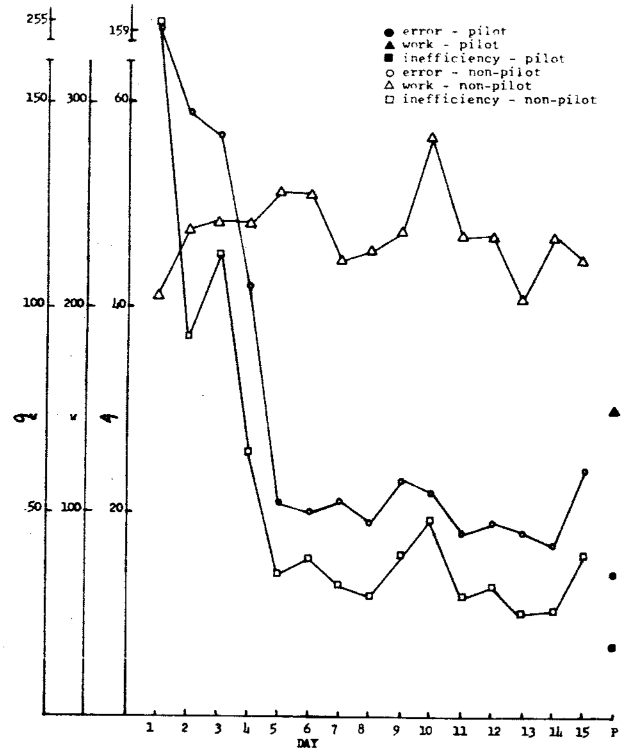

Figure 101 Error (σ) and inefficiency (σw) and work (w) for yaw in the presence of pitch and altitude.

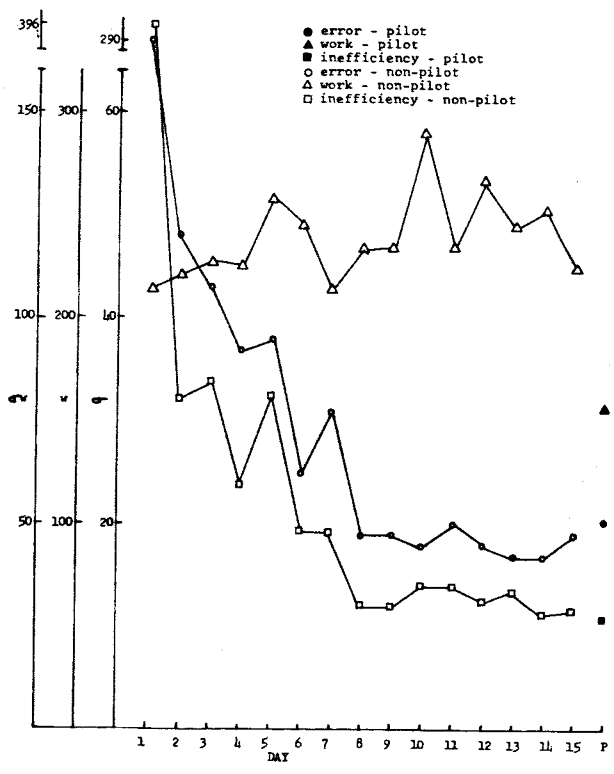

Figure 102 Error (σ) and inefficiency (σw) and work (w) for yaw in the presence of roll and altitude.

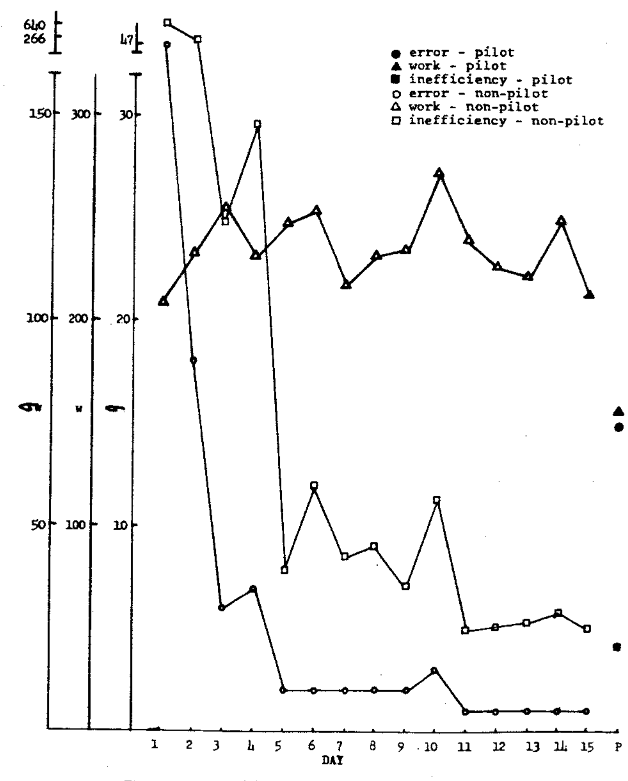

Figure 103 Error (σ) and inefficiency (σw) and work (w) for yaw in the presence of pitch, roll, and altitude.

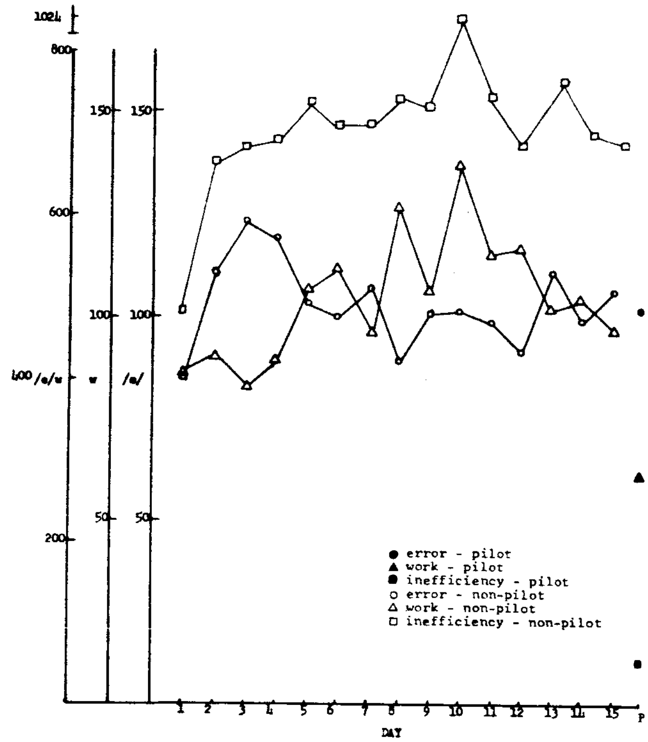

Figure 104 Error /e/ and inefficiency /e/w and work w for altitude alone.

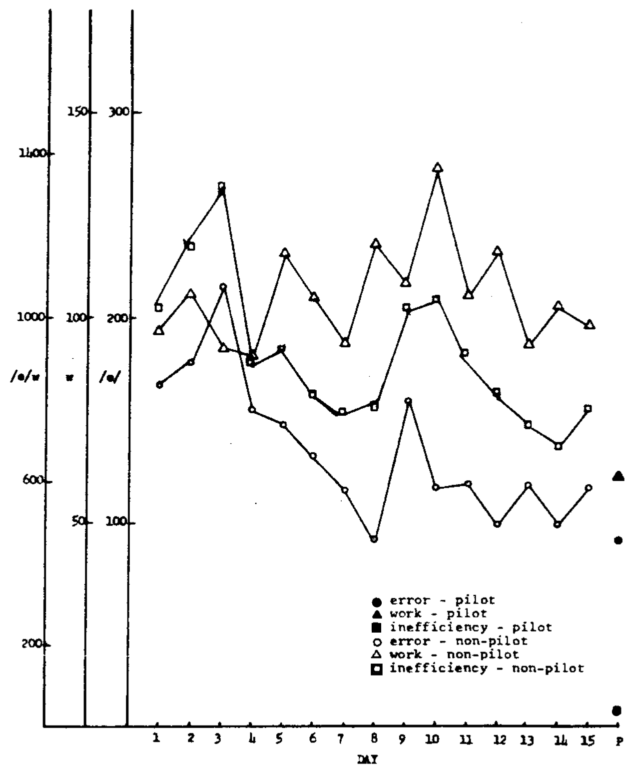

Figure 105 Error (/e/) and inefficiency (/e/w) and work (w) for altitude in the presence of pitch.

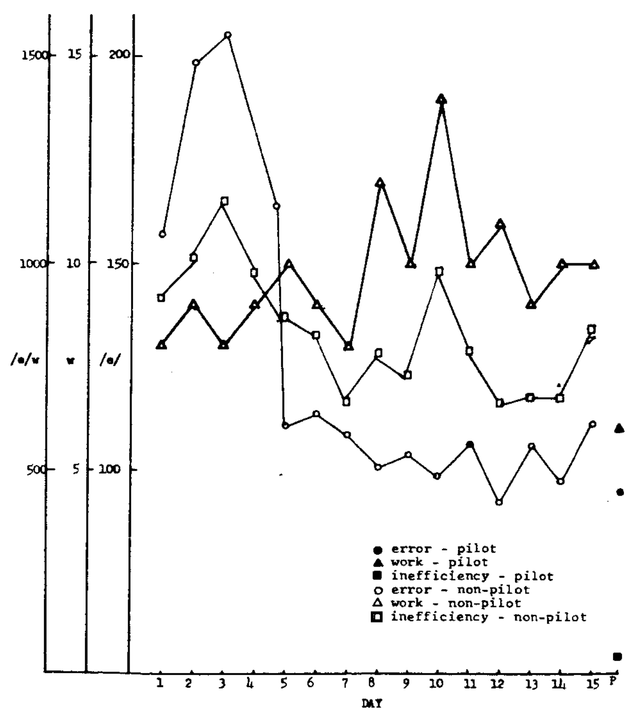

Figure 106 Error (/e/) and inefficiency (/e/w) and work (w) for altitude in the presence of roll.

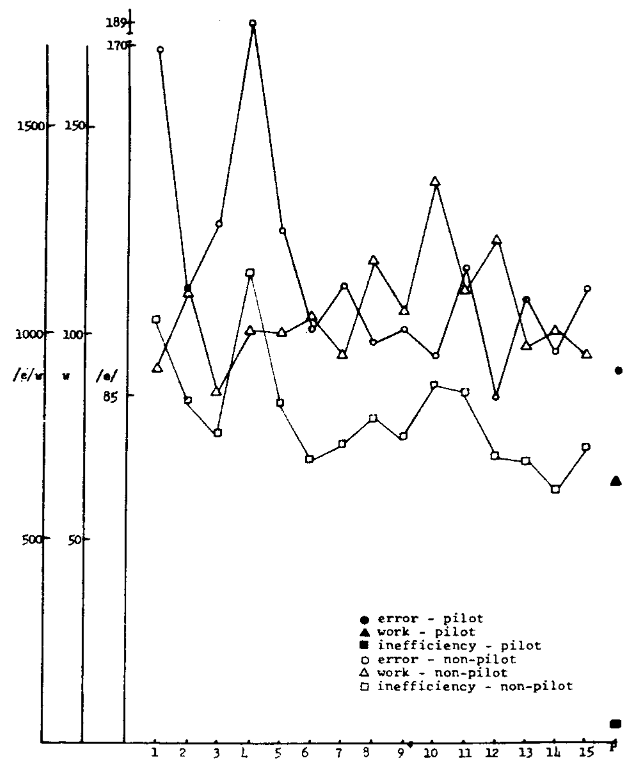

Figure 107 Error (/e/) and inefficiency (/e/w) and work (w) for altitude in the presence of yaw.

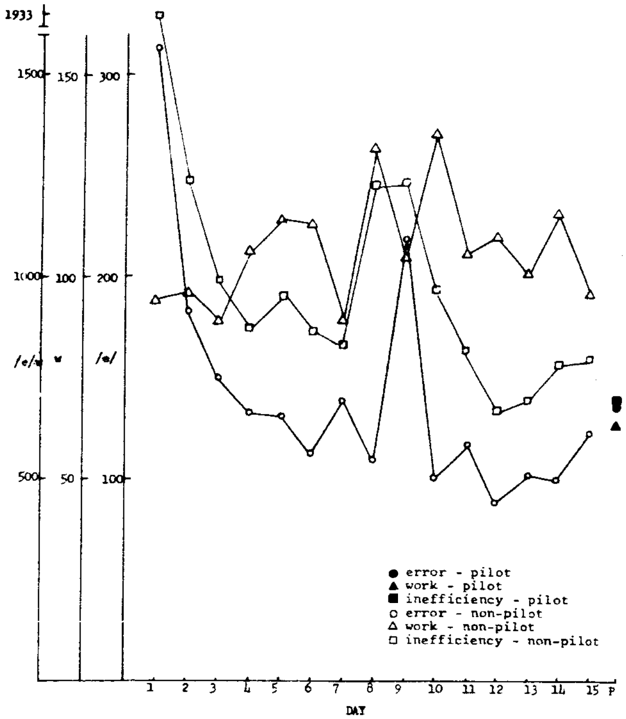

Figure 108 Error (/e/) and inefficiency (/e/w) and work (w) for altitude in the presence of pitch and roll.

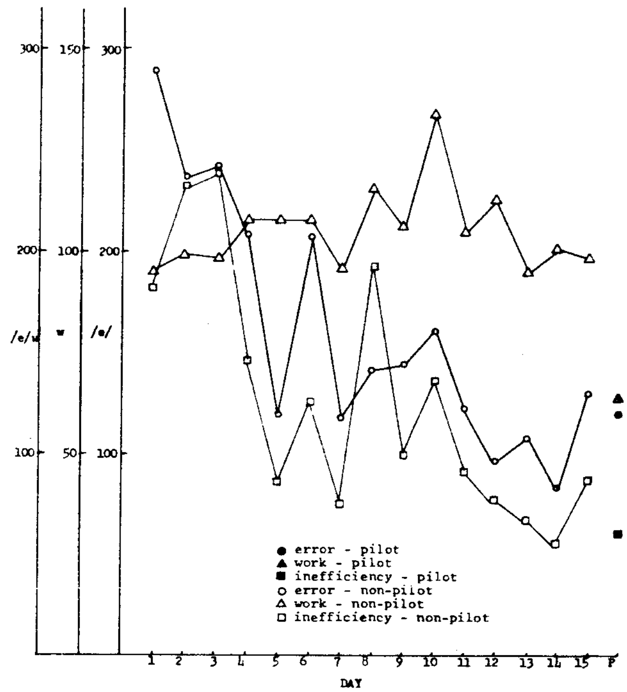

Figure 109 Error (/e/) and inefficiency (/e/w) and work (w) for altitude in the presence of pitch and yaw.

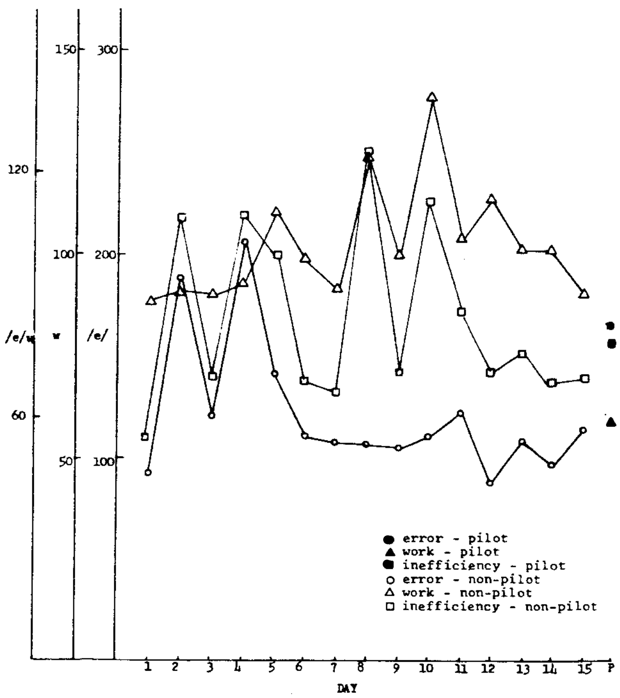

Figure 110 Error (/e/) and inefficiency (/e/w) and work (w) for altitude in the presence of roll and yaw.

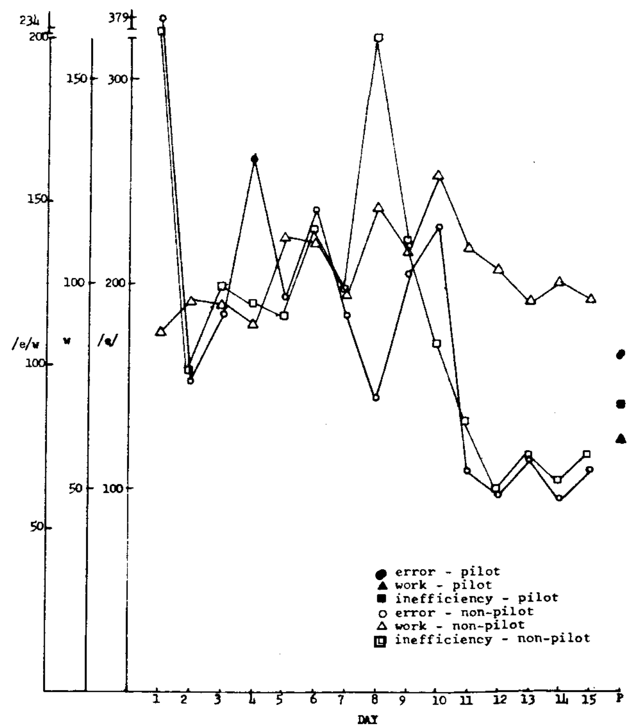

Figure 111 Error (/e/) and inefficiency (/e/w) and work (w) for altitude in the presence of pitch, roll, and yaw.

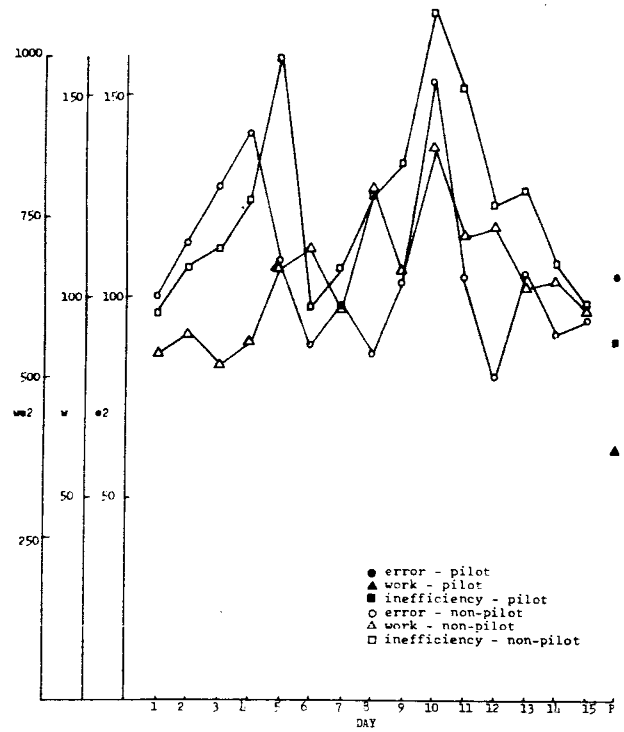

Figure 112 Error (e2) inefficiency (we2) and work (w) for altitude alone.

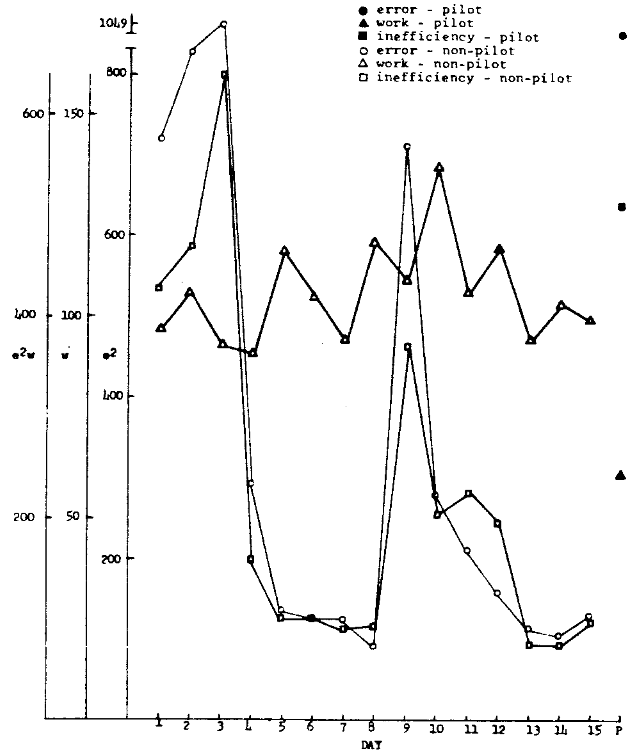

Figure 113 Error (e2) inefficiency (e2w) and work (w) for altitude in the presence of pitch.

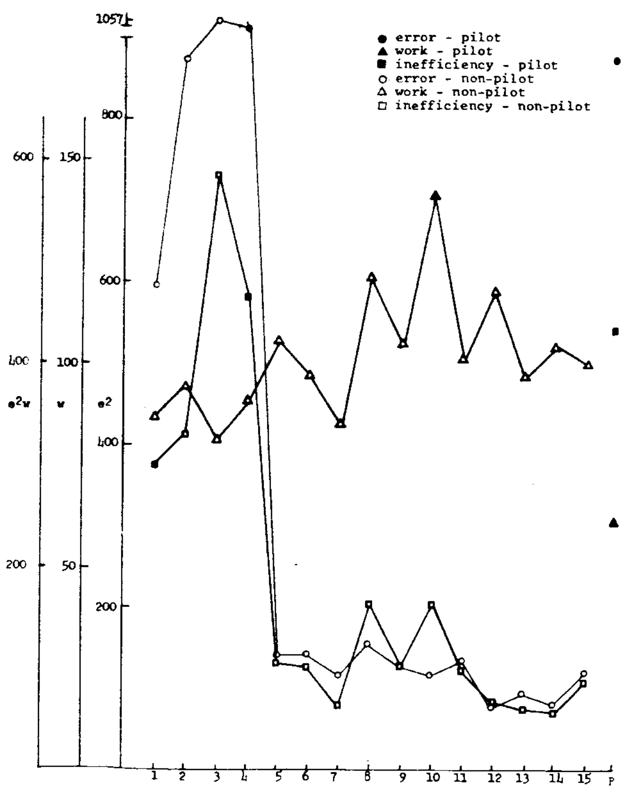

Figure 114 Error (e2) inefficiency (e2w) and work (w) for altitude in the presence of roll.

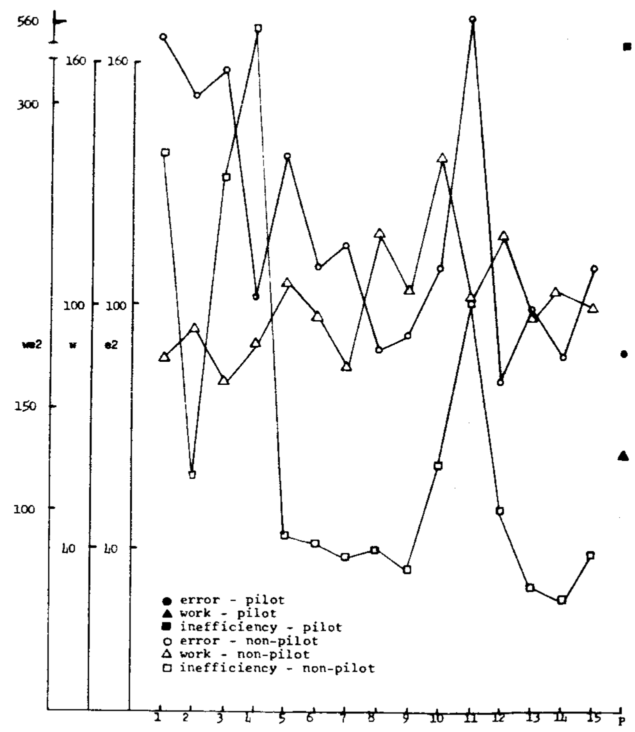

Figure 115 Error (e2) inefficiency (we2) and work (w) for altitude in the presence of yaw.

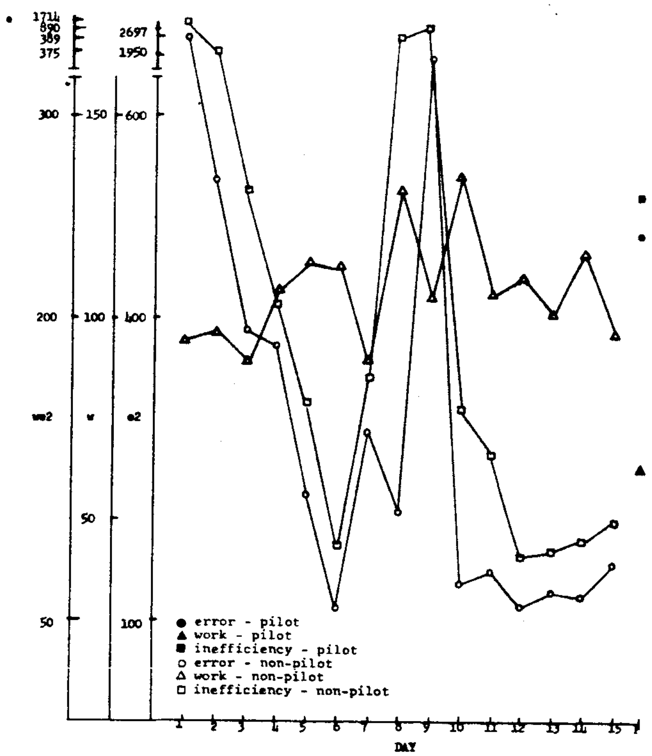

Figure 116 Error (e2) ineffeciency (we2) and work (w) for altitude in the presence of pitch and roll.

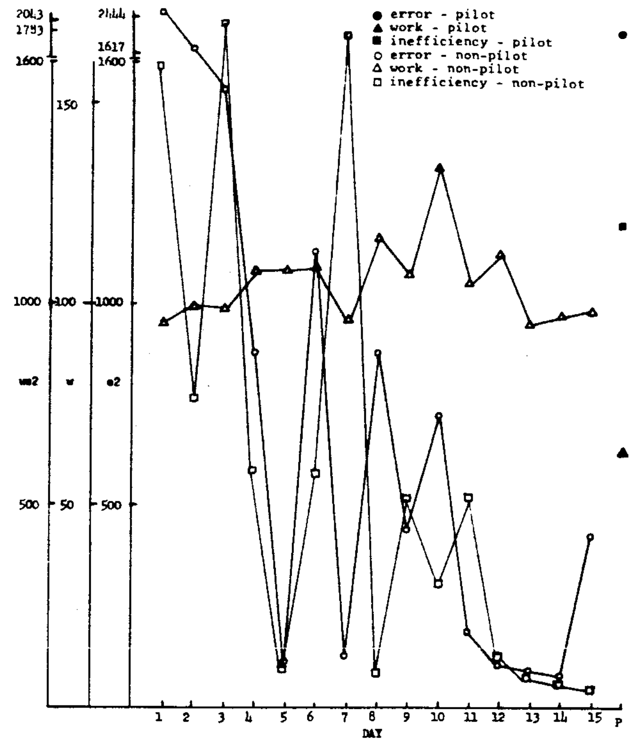

Figure 117 Error (e2) inefficiency (we2) and work (w) for altitude in the presence of pitch and yaw.

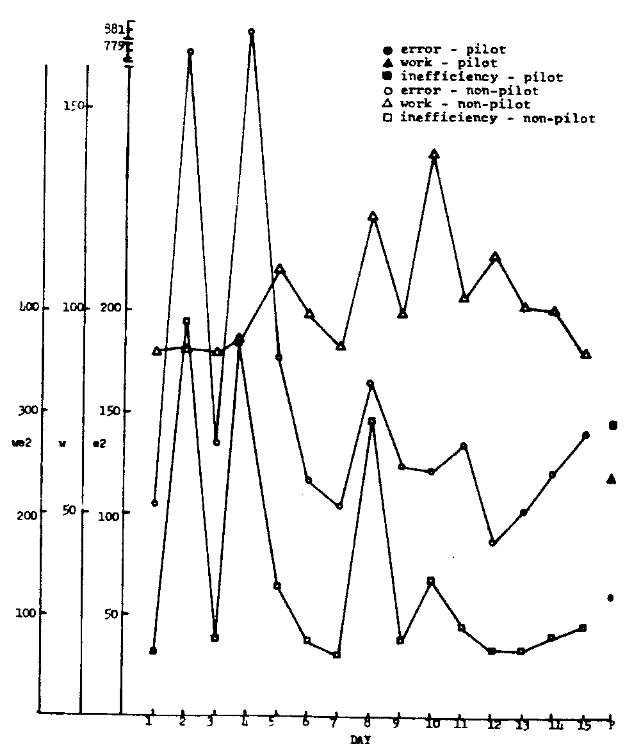

Figure 118 Error (e2) inefficiency (we2) and work (w) for altitude in the presence of roll and yaw.

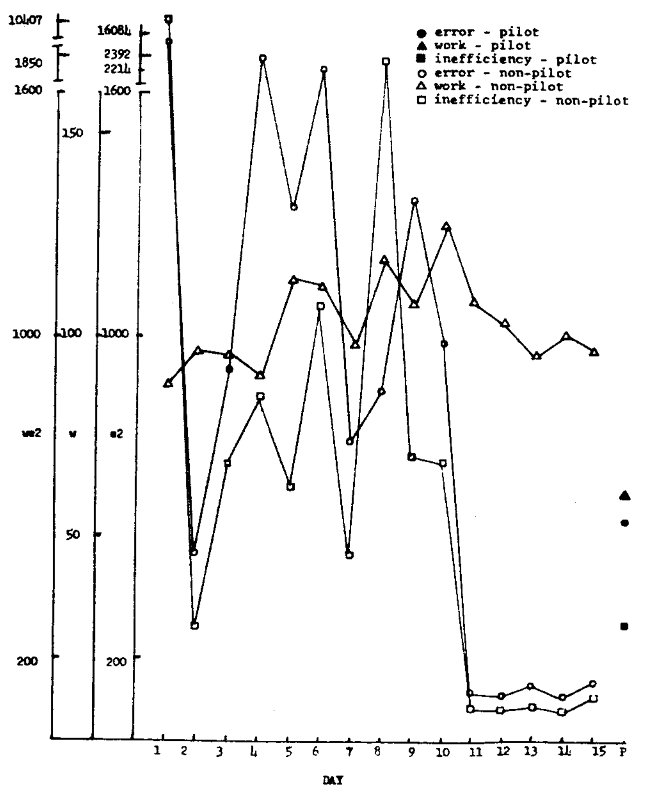

Figure 119 Error (e2) inefficiency (we2) and work (w) for altitude in the presence of pitch, roll and yaw.

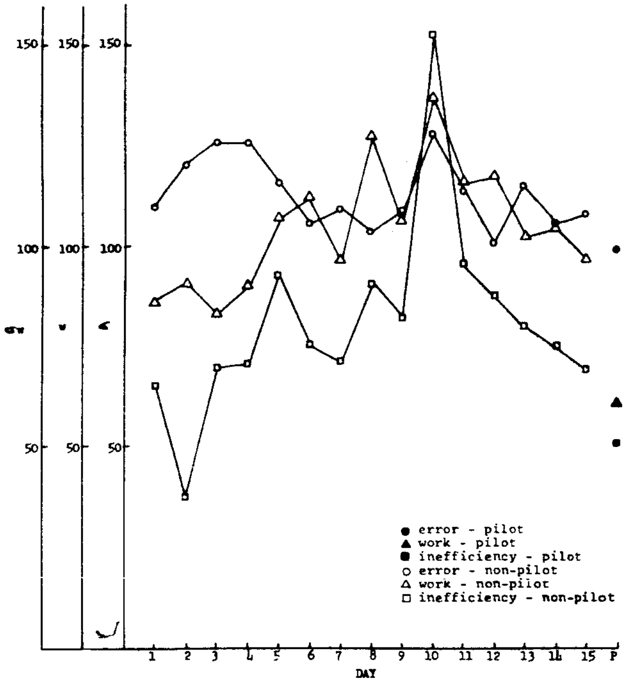

Figure 120 Error (σ) inefficiency (σw) and work (w) for altitude alone.

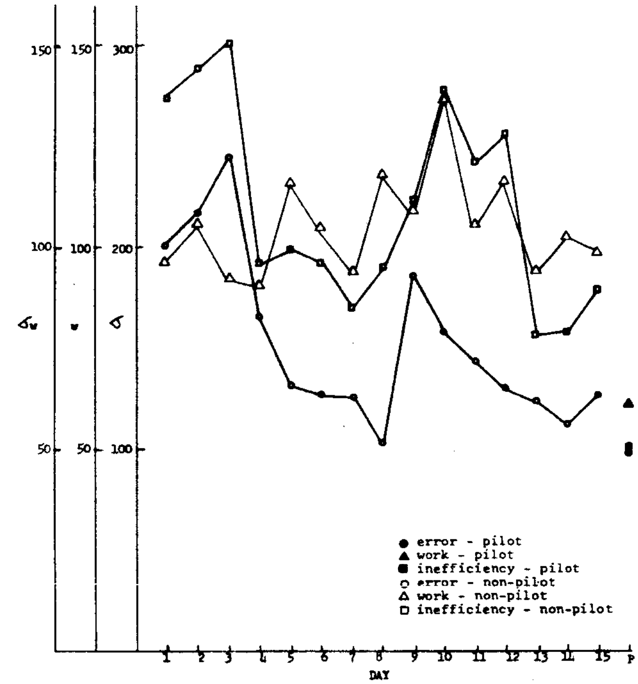

Figure 121 Error (σ) inefficiency (σw) and work (w) for altitude in the presence of pitch.

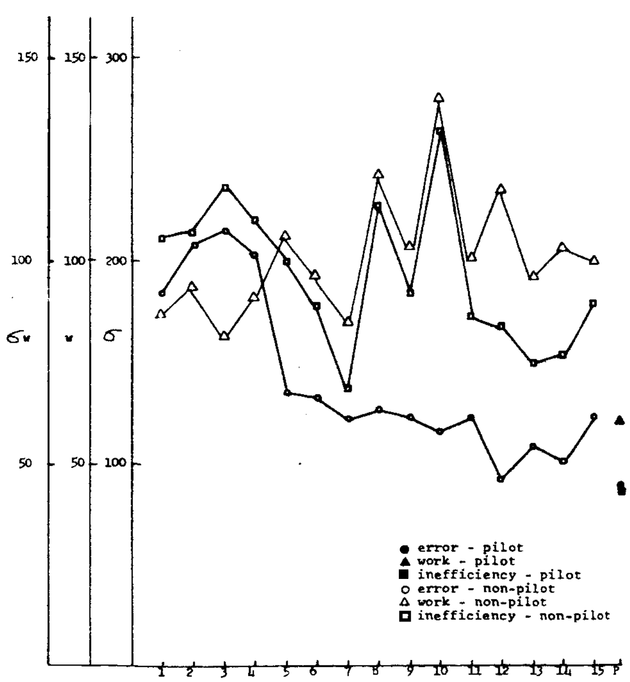

Figure 122 Error (σ) inefficiency (σw) and work (w) for altitude in the presence of roll.

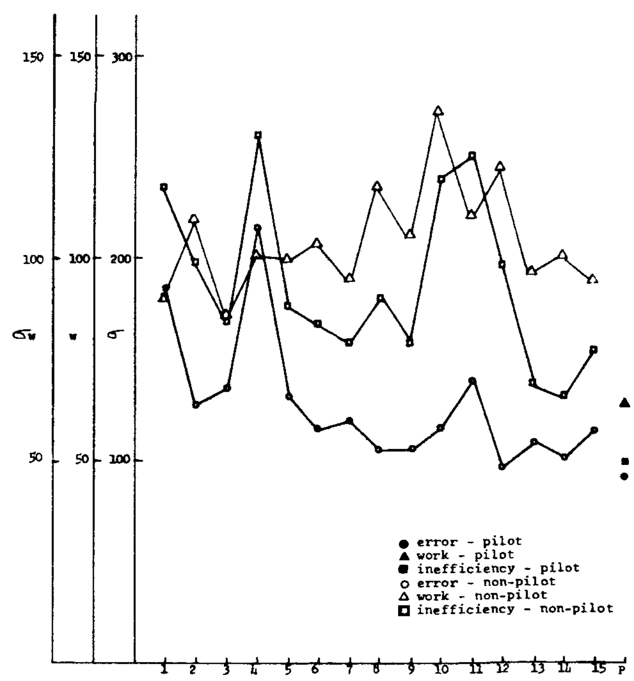

Figure 123 Error (σ) inefficiency (σw) and work (w) for altitude in the presence of yaw.

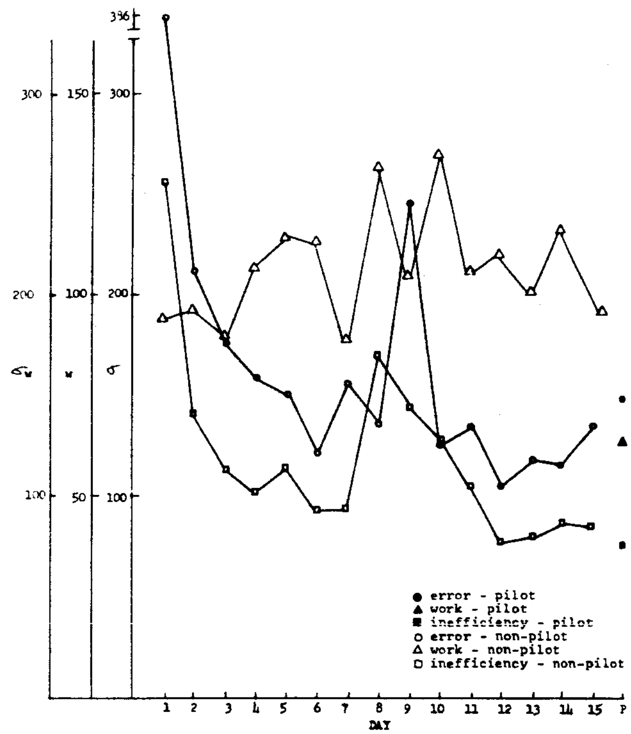

Figure 124 Error (σ) inefficiency (σw) and work (w) for altitude in the presence of pitch and roll.

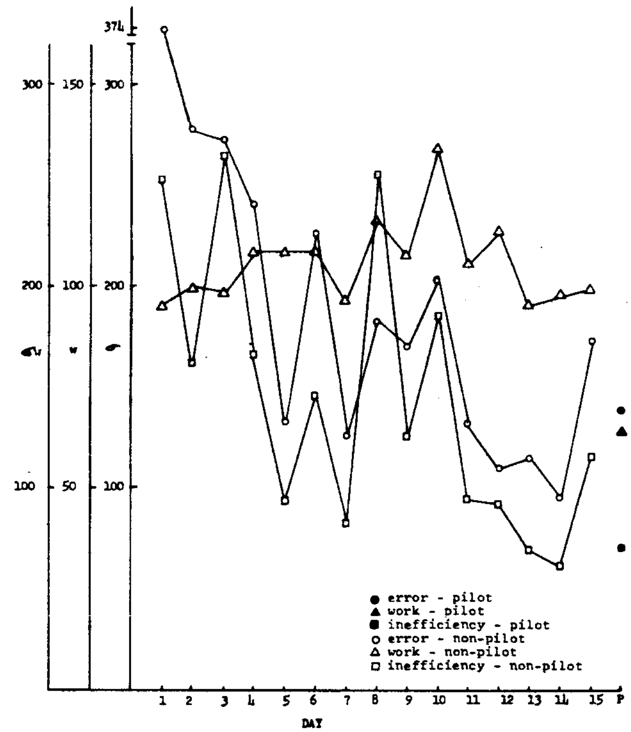

Figure 125 Error (σ) and inefficiency (σw) and work (w) for altitude in the presence of pitch and yaw.

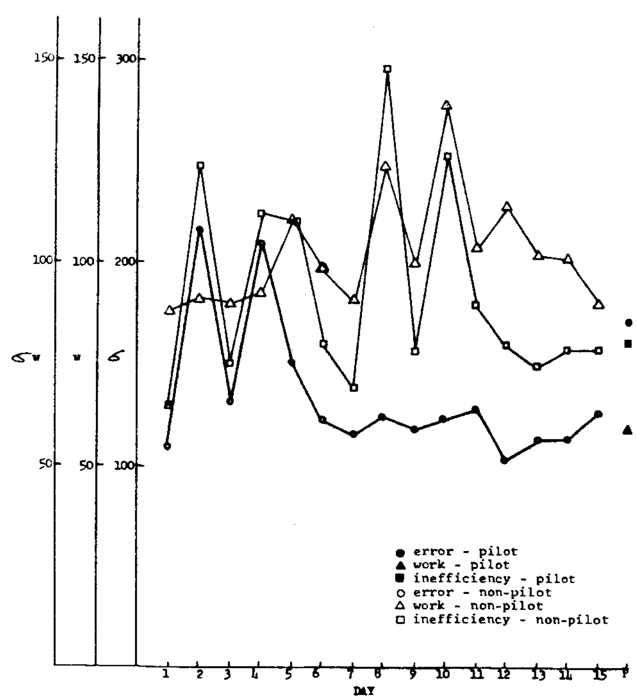

Figure 126 Error (σ) inefficiency (σw) and work (w) for altitude in the presence of roll and yaw.

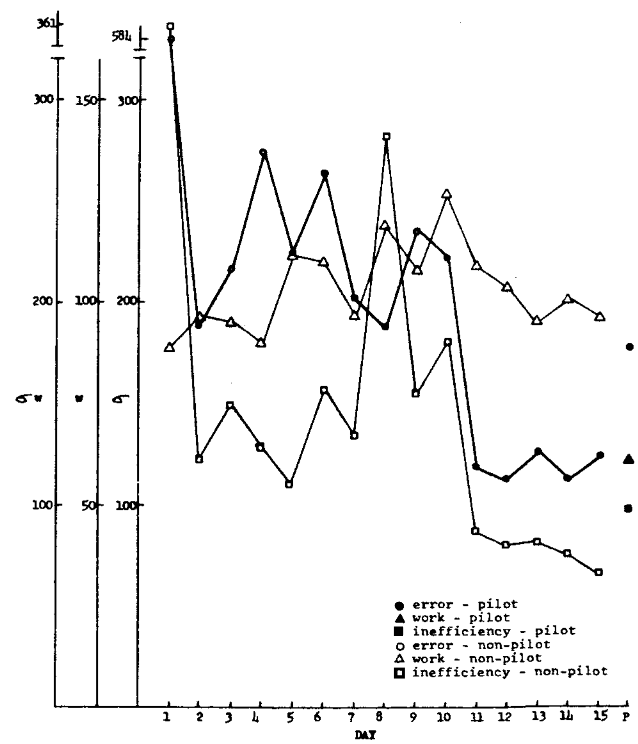

Figure 127 Error (σ) inefficiency (σw) and work (w) for altitude in the presence of pitch, roll and yaw.

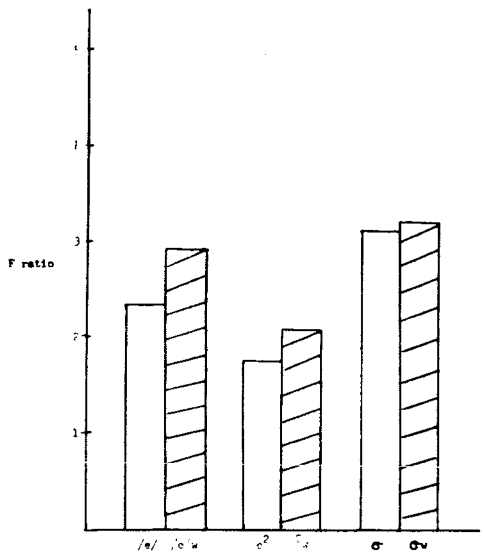

Figure 128 F ratios obtained between conditions for pitch (pilot data).

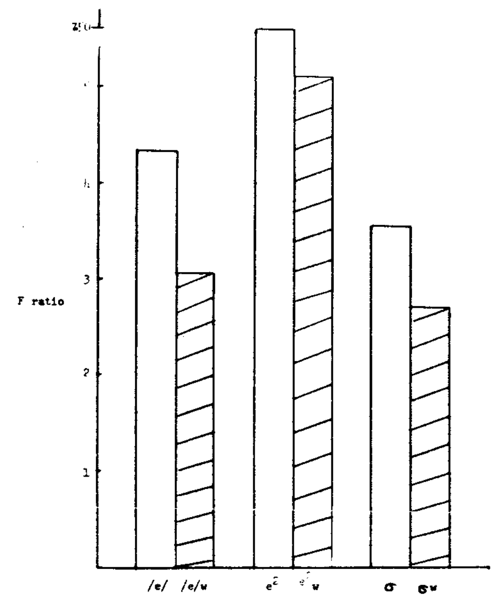

Figure 129 F ratios obtained between subjects for pitch (pilot data).

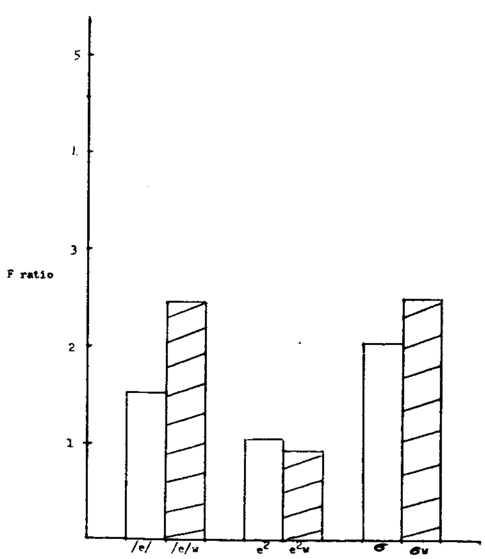

Figure 130 F ratios obtained between conditions for roll (pilot data).

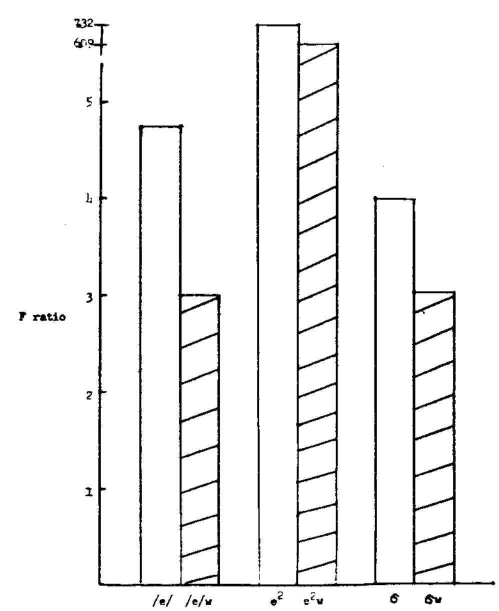

Figure 131 F ratios obtained between subjects for roll (pilot data).

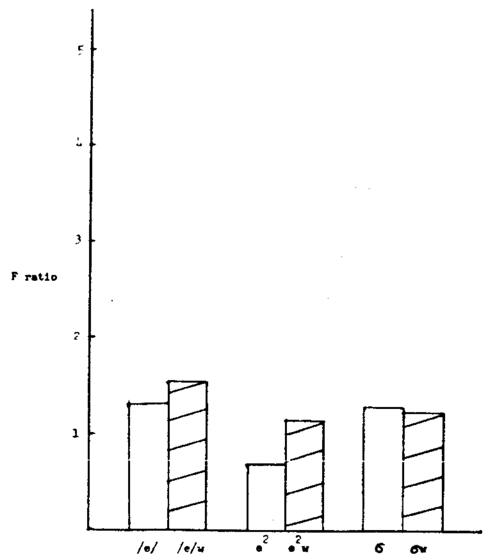

Figure 132 F ratios obtained between conditions for yaw (pilot data).

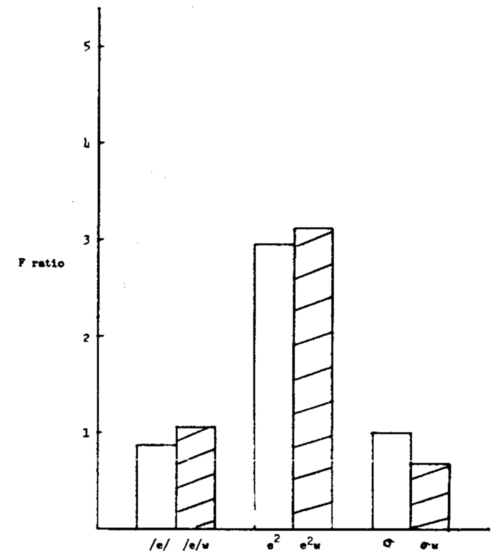

Figure 133 F ratios obtained between subjects for yaw (pilot data).

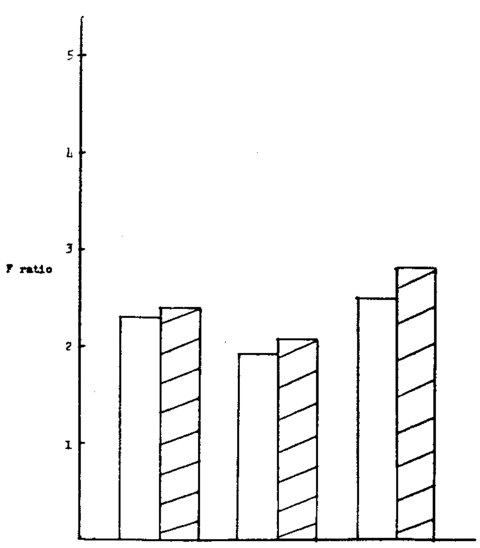

Figure 134 F ratios obtained between conditions for altitude (pilot data).

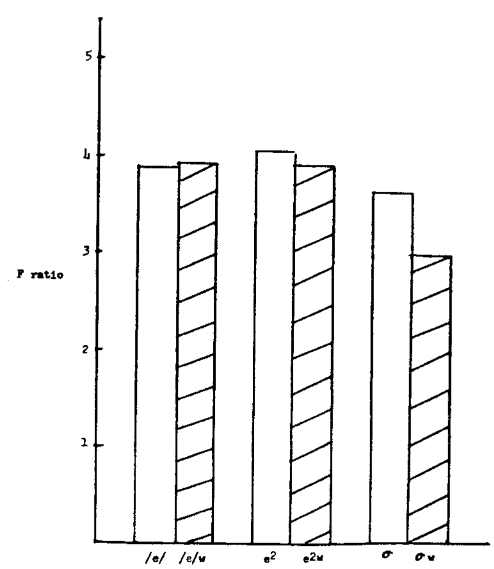

Figure 135 F ratios obtained between subjects for altitude (pilot data).

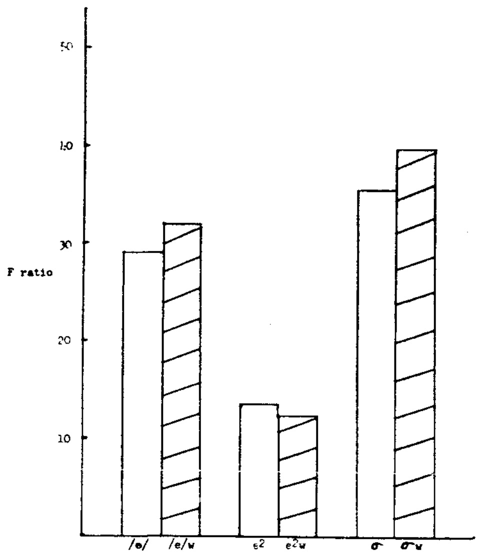

Figure 136 F ratios obtained between conditions for pitch (non-pilot data).

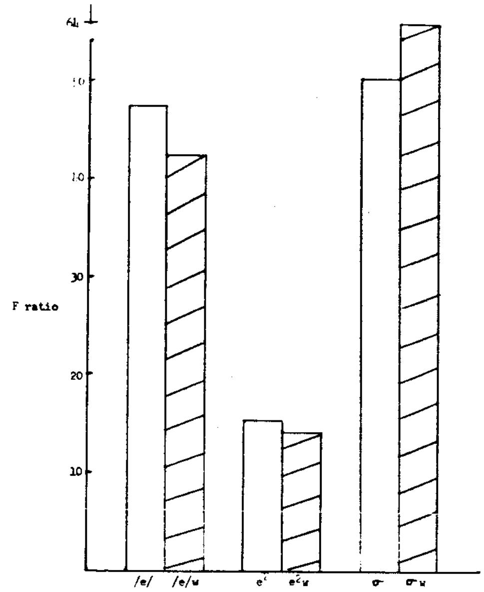

Figure 137 F ratios obtained between days for pitch (non-pilot data).

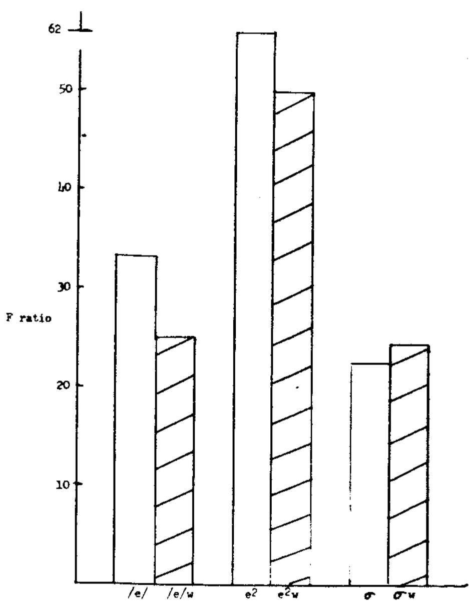

Figure138 F ratios obtained between subjects for pitch (non-pilot data).

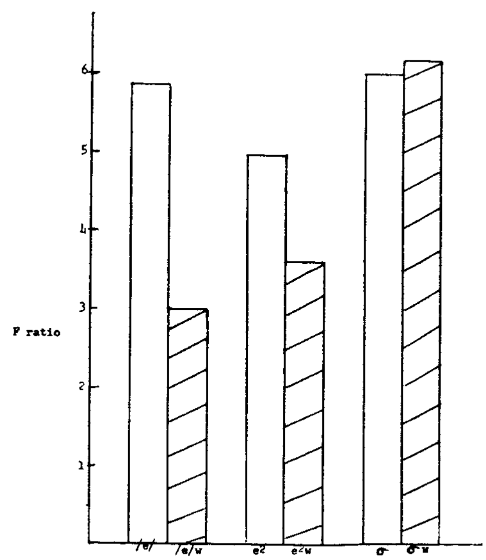

Figure 139 F ratios obtained for conditions x days interaction for pitch (non-pilot data).

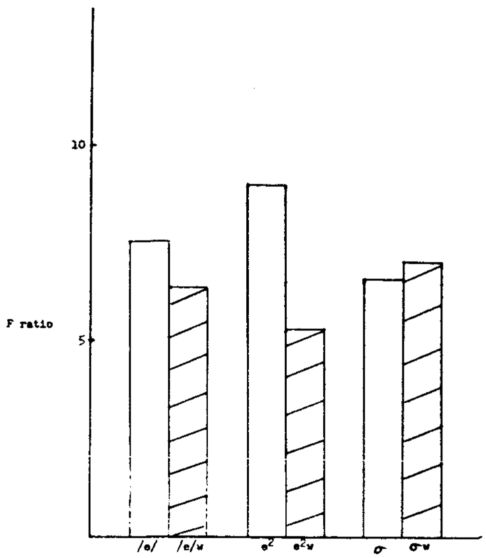

Figure 140 F ratios obtained for conditions x subjects interaction for pitch (non-pilot data).

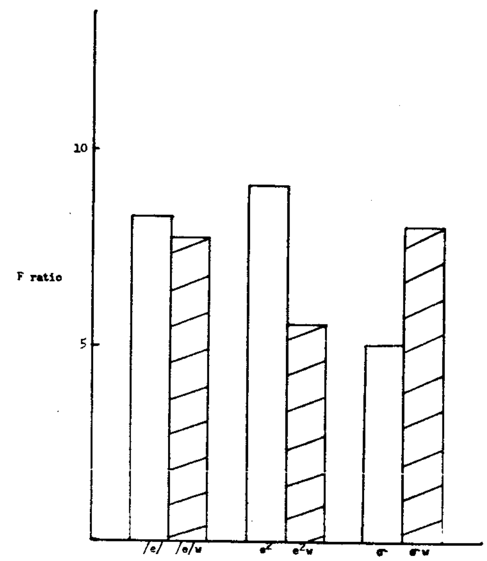

Figure 141 F ratios obtained for days x subjects interaction for pitch (non-pilot data).

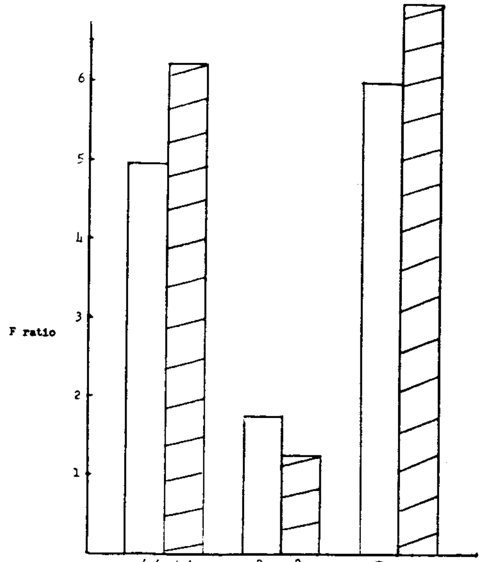

Figure 142 F ratios obtained between days for roll (non-pilot data).

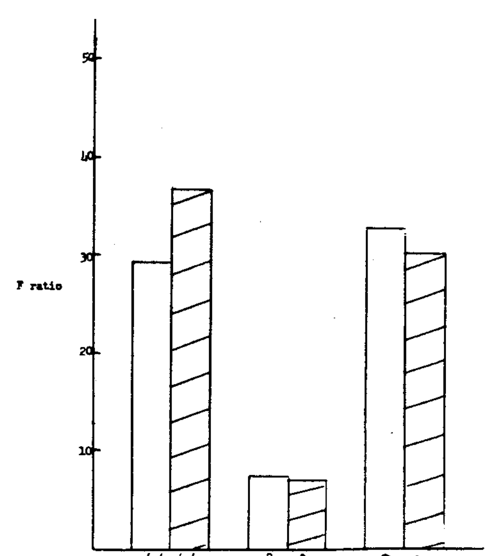

Figure 143 F ratios obtained between conditions for roll (non-pilot data).

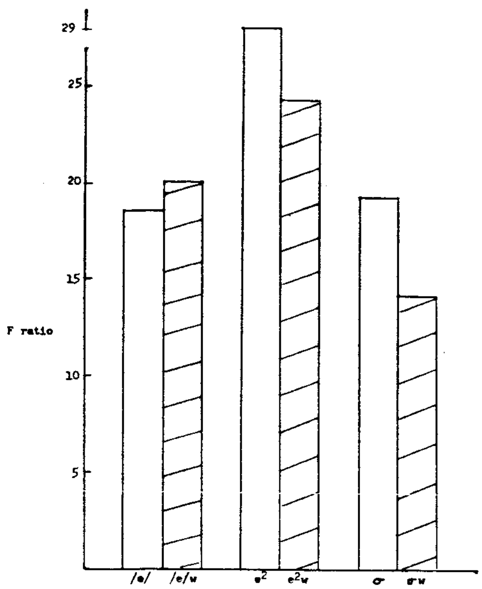

Figure 144 F ratios obtained between subjects for roll (non-pilot data).

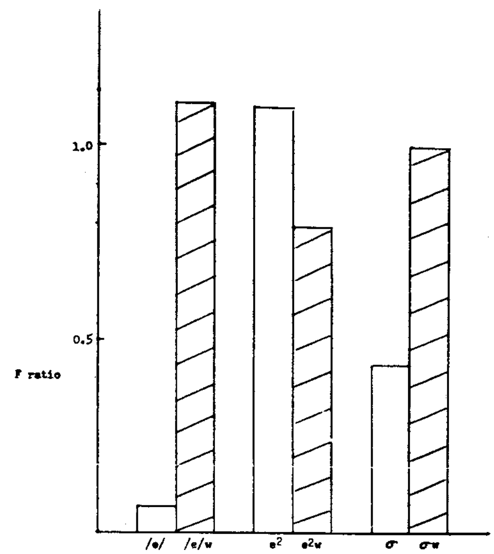

Figure 145 F ratios obtained for conditions x days interaction for roll (non-pilot data).

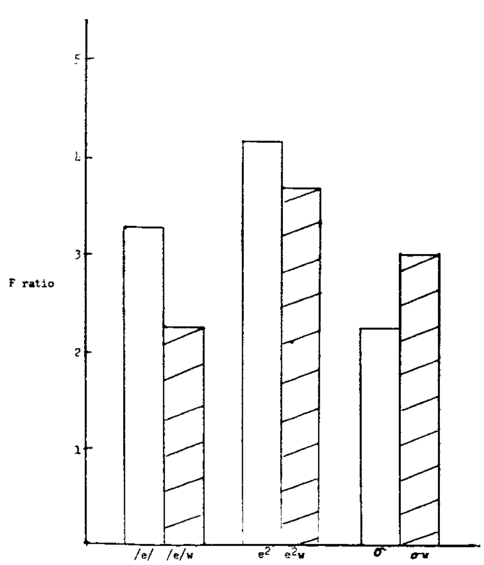

Figure 146 F ratios obtained for conditions x subjects interaction for roll (non-pilot data).

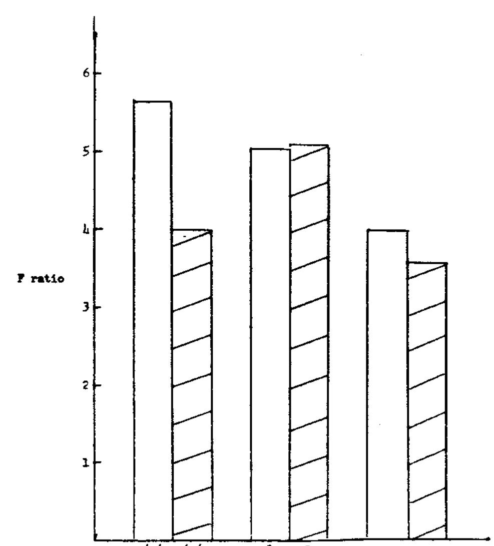

Figure 147 F ratios obtained for days x subjects interaction for roll (non-pilot data).

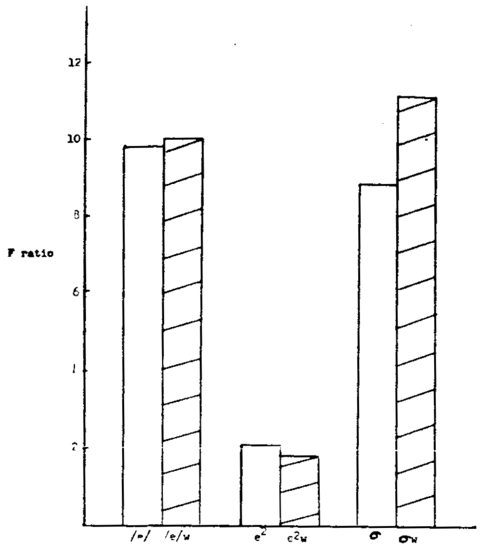

Figure 148 F ratios obtained between conditions for yaw (non-pilot data).

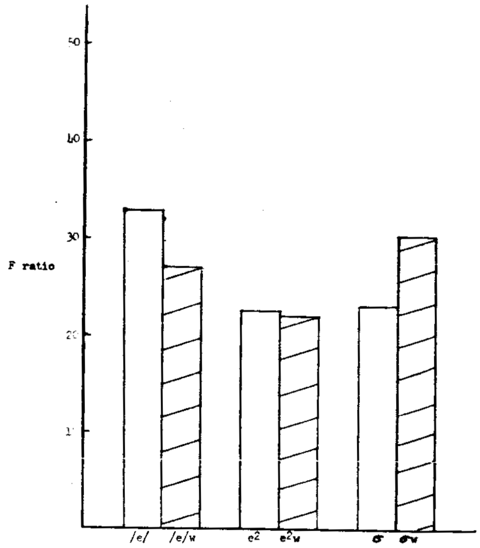

Figure 149 F ratios obtained between days for yaw (non-pilot data).

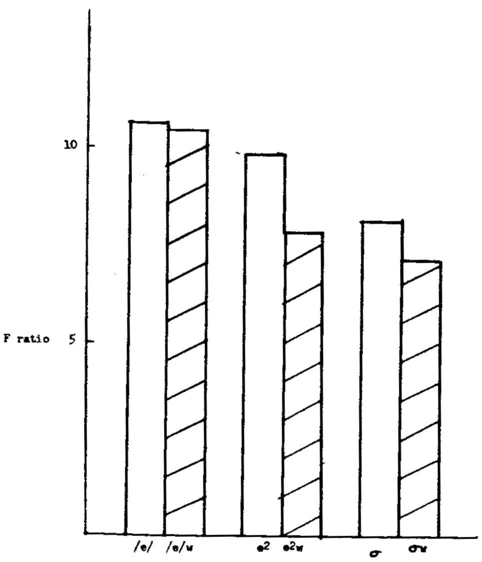

Figure 150 F ratios obtained between subjects for yaw (non-pilot data)

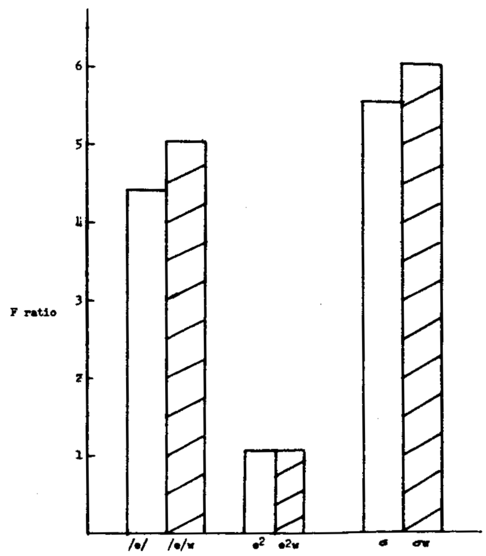

Figure 151 F ratios obtained for conditions x days interaction for yaw (non-pilot data)

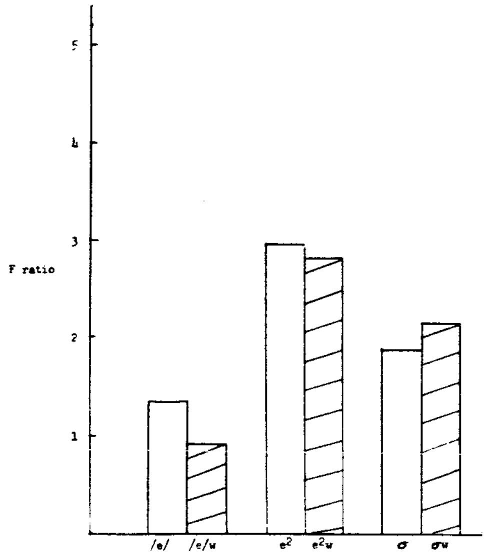

Figure 152 F ratios obtained for conditions x subjects interaction for yaw (non-pilot data).

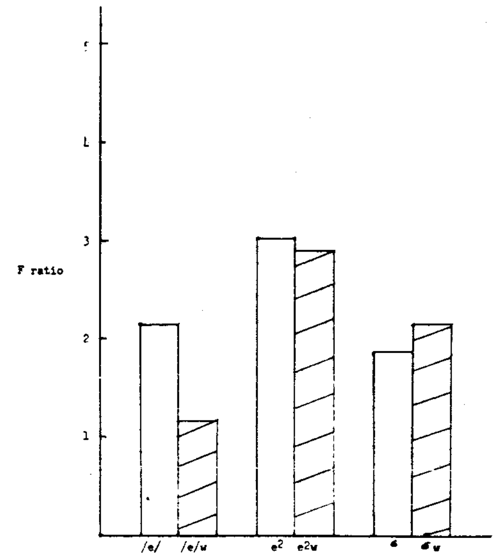

Figure 153 F ratios obtained for days x subjects interaction for yaw (non-pilot data).

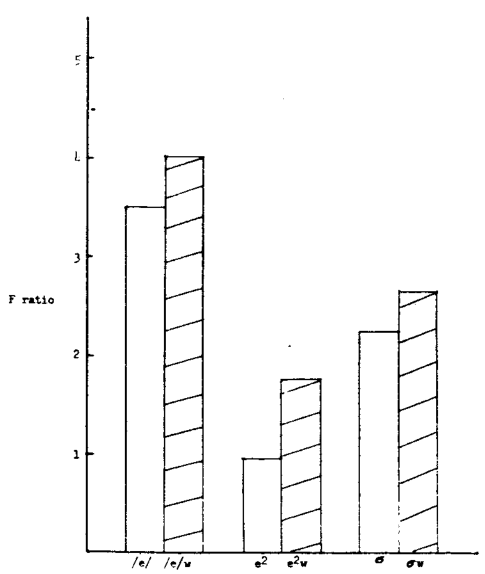

Figure 154 F ratios obtained between conditions for altitude (non-pilot data).

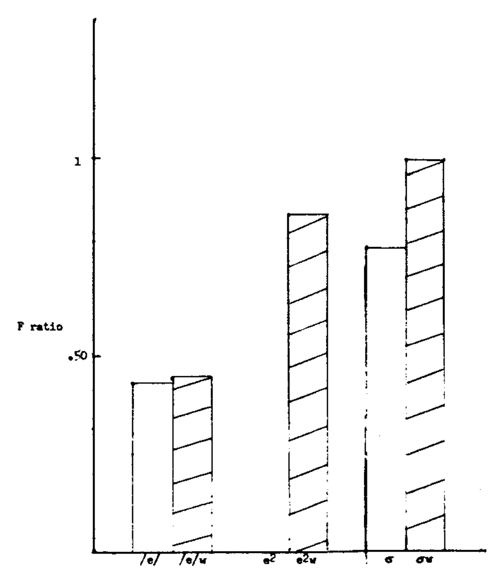

Figure 155 F ratios obtained between days for altitude (non-pilot data).

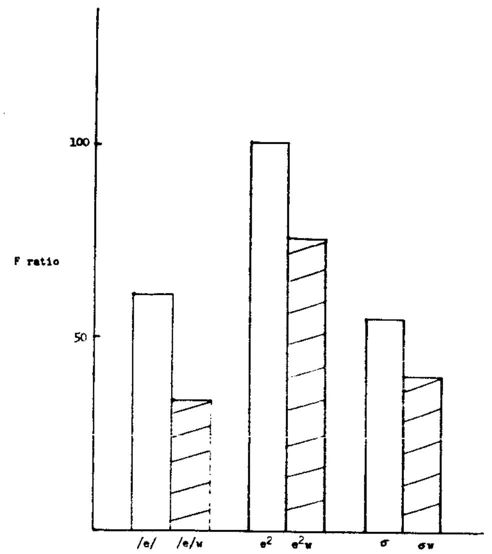

Figure 156 F ratios obtained between subjects for altitude (non-pilot data).

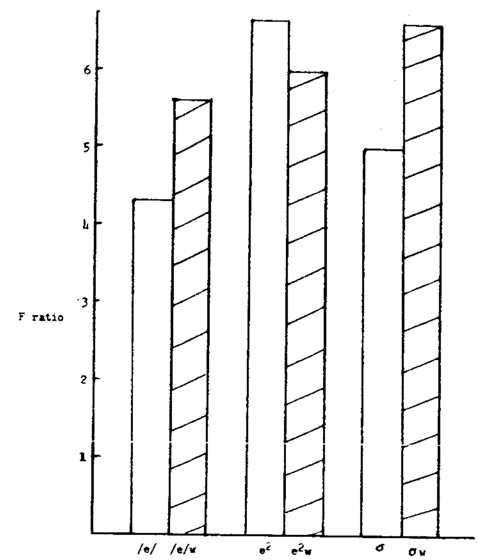

Figure 157 F ratios obtained for conditions x days interaction for altitude (non-pilot) data.

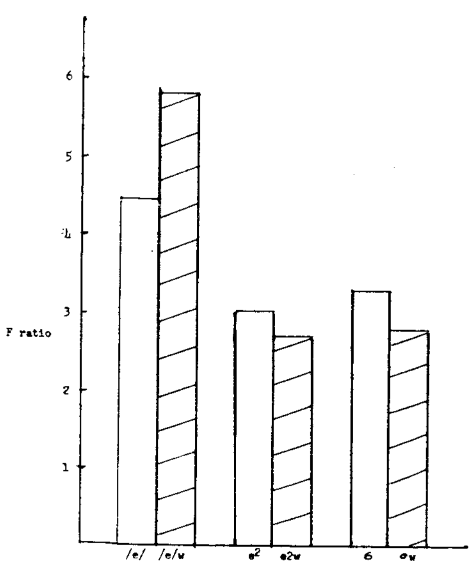

Figure 158 F ratios obtained for conditions x subjects interaction for altitude (non-pilot data)

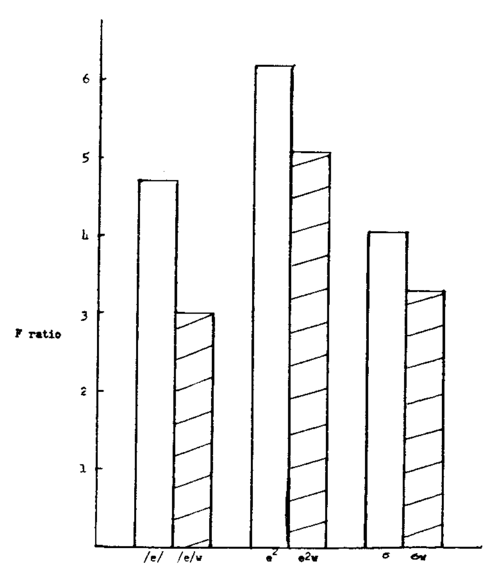

Figure 159 F ratios obtained for days x subjects interaction for altitude (non-pilot data)

VI. DISCUSSION

From the relative scores shown in Figures 4 through 7, it could be hypothesized that with continued training, the proficiency on a task in combination with other tasks approaches the level of proficiency exhibited by the task when given alone. Although the task combinations had not all achieved the same level at the end of fifteen days of training, neither had they become absolutely asymptotic. This hypothesis would obviously be of limited validity and utility, however, since it could not be true for a very large number of tasks, and since the training required before equalization was reached, it could be very great. For a limited number of tasks, however, it could prove a useful rule as a guide in problems of training and design since it would make it unnecessary to test all tasks in combination with one another before judging the eventual proficiency that would be attained on each.

Manifest in these data (see Figures 8 through 31) is that some tasks produce a debilitating interaction with one another while others do not. In the present instance, this is most evident in the case of pitch and altitude control, although a similar synergistic relationship can be found between roll and yaw. This illustrates, as well as anything can, the difficulty of predicting performance as a simple function of the number of tasks. It is unlikely that these instances of conflict can be anticipated prior to testing. In the present case, it is likely that the difficulty is conceptual rather than due to either stimulus or response differentiation. Even non-pilots anticipated a change in pitch and this association was difficult to overcome, although these parameters were independently related in the present experiment.

It is interesting that for differences between treatments the standard deviation proved to be the most sensitive of the error measurements and that the squared error tended to produce the lowest difference among the F ratios. This was largely achieved due to the fact that the between subject variance was lower for the former measurement. This is interpreted to mean that the standard deviation was in closer correspondence with the subject's perception of error than the other measurements were. These effects were common to both pilot and non-pilot subjects. This should not be taken to indicate an unqualified endorsement of the standard deviation as a measure of error, since it is likely that any number of other functions could be contrived which would prove as satisfactory.

Experiments serve many purposes. Infrequently, but occasionally, a study is made in which a novel or unsuspected phenomenon is revealed. More often, experiments establish quantitative values for effects that are known to exist, but which have not been systematically described. Still more frequently, studies merely serve to clarify ones thinking about the problem thus enabling subsequent workers to state the questions in more rational terms.

While the present work may lay some claim to each of these catagories, it should be confessed that its contribution is more one of clarification than of discovery. The original intent was to collect data in such a manner that some mathematical expression could be written in description of performance as it varies with task loading. On the face of it, this appeared to be a straightforward objective requiring scrupulous care in the collection of the data and some sophistication in the development of the mathematical model. When analysis was attempted, however, the basic concepts of the study begin to appear more nebulous than was originally supposed. Without self castigation, it can perhaps be said that a thoughtful experimenter would have recognized this from the beginning. Without undue rationalization, however, it can also be said that the present experimenter was in good scientific company while progressing toward a logical cul de sac.

It now appears doubtful if the notion of "task" has a great deal of utility or validity in purely theoretical research. The difficulties are many. As pointed out earlier, the term is very broad and may include anything from the most simple to the most complex of behavior. In some types of research, this might pose no problem, but when an attempt is made to combine tasks, it is necessary to define the common units of which the tasks are constituted. Some approach couched in stimulus response or signal detection language would probably be futile and certainly laborious and distasteful. Added to this there is a species of circularity in the concept. In a way that is peculiar to itself, the idea of "task" cannot exist outside our method of measuring it. Some behavior can be described without reference to values or to volitional elements but the notion of "task" clearly implies purpose as well as judgment with respect to a criteria. This would not be troublesome were it not for the fact that the criteria established by the experimenter's analysis is almost certainly different from that accepted by the subject. It is not very useful to speak of performance if we mean by that term something other than the difference between the subject's behavior and how he intends to behave. On the other hand, if every subject is allowed to produce his own criteria, we are faced with the anarchy of relativism that is euphemistically referred to as pilot judgment.

In the opinion of the present experimenter, there is no satisfactory resolution to this dilemma within the context of "pure" research. The problem becomes real only when the criteria applied is pragmatically well grounded. This is to say that it must be realistically based upon such obvious things as safety of flight, probability of mission success, etc. These values must be recognized by the subject as clearly as they are by the experimenter. Experimental conclusions based upon such measures as absolute error and mean square error have little place in an applied problem. The error scale should be tailored to fit the task.

From the present analysis there is evidence that the inefficiency indices are of greater value in the assessment of performance than are the error scores. This is plausible considering the fact that they are the product of error and subject activity. In a sense, the subject evaluates himself by the amount of work he is doing. It will be recognized, however, that these are rather crude estimates of inefficiency since they only reflect rate of control movement which is not especially well correlated with work performed. One might be led to conclude that acceleration of control movements would prove to be the best reference to subject activity. From a purely mechanical viewpoint, this is, of course, correct but more than simple mechanics is involved. A subject may report being exhausted after an interval in a tracking situation despite the fact that the physical work performed was negligible. It seems evident that work in this case relates more to the perceptual-cognitive-emotional processes than it does to muscular activity although some of each may be involved. Let us suppose we call this complex "psychological work." We are now left with the question of how we go about measuring it. The most manifest thing about it is that it produces a high rate of decision making. Even in a simple tracking task, the subject makes decisions at a rapid rate. The result of each decision is to make an input of a quantity that may vary from zero to the limit of anatomical-physiological tolerance. The force of the input would not, however, correlate particularly well with the notion of psychological work. An input, whatever its magnitude, is unitary, being the response to a single decision.

A better measurement would appear to be the rate of change of acceleration. Assuming that the coefficient of friction does not vary with the rate of control movement, it can be said that changes in acceleration indicate changes in perception or decision. A measurement of this parameter should consequently provide the best measurement of psychological work. The sum of the products of the momentary values of the rate of change of acceleration and the pragmatic error score should produce a reasonable estimate of inefficiency.

VII. SUMMARY

An experiment was conducted in which four (roll, pitch, yaw and altitude control) tracking tasks were presented in all possible combinations with one another. The intent of the study was to contribute to an understanding of how performance is affected as the task loading is increased. Fifteen pilots and fifteen non-pilot subjects were used. A helicopter simulator mounted on a dynamic platform was used as the principle apparatus of the study. Information concerning the control parameters was transmitted to the subject via the Norden vertical contact analog display. Momentary error and control position were recorded on magnetic tape.

Six basic manipulations were performed on the data. Results were presented in terms of absolute error, squared error and the standard deviation about the mean error. In addition, indices of subject inefficiency were produced by obtaining the products of each error with the rate of control movement.

The following conclusions are to be drawn from the data:

1. With training, proficiency on a task in the presence of other tasks approached the proficiency exhibited on the task presented alone.

2. Although error decreased with experience in a negatively accelerated manner, the amount of work accomplished in the performance of the task did not show a concomitant decrease.

3. Certain tasks (pitch and altitude control will serve as an example) when combined tended to interfere with one another. Others did not show this interference. This indicates that it is impossible to predict performance as a simple function of the number of tasks that are combined. In the present instance, this interference appears to be due more to perceptual than to manual control difficulties.

4. Among the error scores the standard deviation produced the most significant difference between conditions of measurement. The squared error produced the least. As these measures related to subject differences, however, this order was reversed.

5. The indices of inefficiency produced a more significant difference between experimental conditions than did error scores. This again was reversed with reference to subject differences.

6. The results of the tracking study were greatly affected by the scale of measurement used. It was concluded therefore that the scale chosen should

conform realistically to the safety and utility factors of the vehicle simulated.

VIII. REFERENCES

1. Baker, Katherine E., Wylie, Ruth C., and Gagne, R. M. The effects of an interfering task on the learning of a complex motor skill, J. Exp. Psychol., 1951, 41, 1-9.

2. Conrad, R., Some effects on performance of changes in perceptual load, J. Exp. Psychol., 1955, 49, 313-322.

3. Elam, C. B. and Emery, J. Effect of stimulus ambiguity in the display of attitude information, ANIP D228-421-013, September, 1962.

4. Elam, C. B. and Emery, J. Redundancy in the display at spatial orientation, ANIP D228-421-009, August, 1961.

5. McGeoch, J. A., and Irion, A. R. The psychology of human learning (2nd Ed) New York, Longsman, Green, 1952.

6. Jeantheau, G. The differential effects of speed and load stress on task performance. WADC TR 59-7, July, 1959, AD 233-460.

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Bell Helicopter Company P. O. Box 482, Fort Worth, Texas	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP N/A

3. REPORT TITLE
RESEARCH ON UTILIZATION OF PART TASK SPATIAL ORIENTATION INFORMATION IN THE DYNAMIC SIMULATOR

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
NASA Technical Report

5. AUTHOR(S) (Last name, first name, initial)
Elam, C.B.
Abbott, Billie A.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
June 1965	71	6

8a. CONTRACT OR GRANT NO. NASw-439 b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) No. 299-099-284 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) None

10. AVAILABILITY/LIMITATION NOTICES

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
None	National Aeronautics & Space Admin. Washington, D.C.

13. ABSTRACT

Fifteen professional pilots and a like number of R.O.T.C. students were used as subjects in establishing the relative difficulty in controlling for pitch, roll, yaw and altitude when these parameters were presented in all possible combinations with one another. The Bell simulator was used with the Norden vertical display serving as the media for information transmission. Momentary error and control position were recorded and were later converted into error and inefficiency scores.
While highly significant in a statistical sense, the differences between conditions were considered to be of limited therοretical interest. The primary contribution of the study was in its comparison of the different indices of error (absolute, squared and standard deviation) and the measures of inefficiency.
It was concluded that the standard deviation is the most sensitive of the error measurements and that an index of inefficiency (product of the error and the rate of control movement) is more sensitive of error alone.

DD FORM 1 JAN 64 1473

Unclassified
Security Classification